WITTGENSTEIN, RULES AND INSTITUTIONS

WITTGENSTEIN, RULES AND INSTITUTIONS

David Bloor

London and New York

First published 1997
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Typeset in Times by Routledge
Printed and bound in Great Britain by
Creative Print and Design (Wales), Ebbw Vale

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data
Bloor, David.
Wittgenstein on rules and institutions/David Bloor.
Includes bibliographical references and index.
1. Wittgenstein, Ludwig, 1889-1951. 2. Rules (Philosophy) 3. Institutions (Philosophy) 4. Collectivism. 5. Knowledge, Sociology of. I. Title.
B3376.W564B6 1997
192--dc21
97-1475
CIP

ISBN 0–415–16147–9

The motto here is always:
Take a wider look round.
(RFM II: 6)

CONTENTS

PREFACE

There are two competing schools of thought about the nature of rule following. One is usually called 'individualism', the other 'collectivism'. For the individualist, a rule in its simplest form is just a standing intention; for the collectivist, it is a shared convention or a social institution. There are also two competing schools of thought about the character of Wittgenstein's later philosophy and, in particular, his account of rule following. Adherents to one of these schools read him as an individualist; adherents to the other, as a collectivist. Clearly these two issues – the nature of rules and the interpretation of Wittgenstein – are logically independent. No one tries to infer the correct account of rules from Wittgenstein's stance, just as no one decides what Wittgenstein 'must' have thought about rules from a prior decision about their correct analysis. Nevertheless, there is a tendency for these things to go together. Typically, but not always, those who see Wittgenstein as an individualist argue for an individualist account of rule following, while those who see him as a collectivist say collectivism provides the correct account. The argument of this book is no exception. I shall defend a collectivist account of rules, and a collectivist reading of Wittgenstein. Why treat these two issues together? The answer is: because struggling with Wittgenstein's account of rule following, incomplete and fragmentary as it is, is still amongst the best ways to get to grips with the phenomenon itself. To derive maximum benefit from this exercise, of course, we have to get his approach into proper focus. I fear this valuable resource is being taken away from us by the increasingly influential individualist readings, or misreadings, of his work.

I should say at the outset that Wittgenstein did not succeed in refuting individualism. He had no knock-down disproof, and nor

has anyone else. This should occasion no surprise. Under pressure individualism can be turned into a metaphysical claim. (And the same applies to collectivism.) As a piece of metaphysics it can be defended dogmatically by insisting on certain definitions, classifications, dominant metaphors and background assumptions. Fortunately it is now widely accepted that the irrefutability of a metaphysical position is no great intellectual virtue, rendering it immune neither to criticism nor rational comparison. We can measure the intellectual strengths of individualism and collectivism by their relative power to illuminate agreed matters of fact about rules, and agreed and central cases of rule following. In a word: we can assess them *methodologically*. It is in the pursuit of this aim that we can employ Wittgenstein's work to maximum effect.

Some of the philosophical literature on rule following is decidedly ungenerous in its tone. To what extent this is because the themes involved, i.e. individualism and collectivism, have identifiable political resonances, I do not know. In the circumstances, it may be appropriate to say that my defence of a collectivist approach to rules does not derive from any dislike of individualism as such. I do not, in fact, believe there is such a thing as 'individualism as such'. Like collectivism, individualism has no 'essence'. They are both 'family resemblance' phenomena which change over time, taking on a different character and moral significance in different historical circumstances. There have been noble manifestations of individualism as well as despicable ones . Similarly, collectivism has had vile expressions, as well as being the embodiment of admirable moral values. These themes, however, for all their importance, are not the themes of this book, though it is impossible to keep the discussion entirely separate from ideological polemics. My concern is with far more limited questions, but ones sufficiently intriguing to interest analysts of a variety of persuasions, namely: what is a rule, and what is it to follow a rule?

One consequence of the family resemblance character of the categories 'individualism' and 'collectivism' is that the ideological oppositions historically associated with them do not always map neatly on to intellectual oppositions. There are sometimes points of contact and agreement between their advocates where one might at first expect only theoretical divergence. Some insights are shared, and some that are not can be transferred from one tradition and reconstructed in the terms of the other. In this connection it is important to distinguish a generalised ideological individualism

from the approaches to rule following which also fall under that heading. Those who embrace individualism as an ideology are saying how they want society to be organised, e.g. by extending the scope of the market. Individualism, in this sense, is a social form, standing in contrast, perhaps, to more traditional and customary forms. As such, it cannot be at odds with a sociological perspective, because it is just one social arrangement amongst others. The cult of the individual is the product of society. It will fall within the scope of any appropriately general theory, as is clear from such classic work as Tönnies' on *Gemeinschaft* and *Gesellschaft* (Tönnies 1887) and Durkheim's on mechanical and organic solidarity (Durkheim 1893). Advocating individualism as a preferred social form is consistent with accepting that some phenomena may exist, and only exist, as social phenomena. An ideological individualist could, therefore, in principle accept the social character of rule following, that is, reject an individualistic analysis of rule following. This might involve some loss of face, and some danger of terminological confusion, but there need be no logical inconsistency. I have not scrupled to exploit this fact in making comparisons between the collectivist Wittgenstein and thinkers who, for the most part, would be intuitively placed in the opposing tradition.

One further terminological point. In Chapter 3 I shall be defending and using a doctrine G. E. M. Anscombe has, for better or for worse, called 'linguistic idealism'. Since I shall be conjoining this with a sociological approach some critics of sociology will feel their worst fears have been confirmed. 'At last, a sociologist who admits the truth – they are idealists, they don't believe in facts or truth or objective reality.' No one, however, who takes care to understand the meaning of the words 'linguistic idealism' could draw such conclusions. Linguistic idealism, as Anscombe defines it, is consistent with a robust realism about material objects, physical processes, or organic nature. The only 'reality' threatened by it is the mythical reality attributed to 'abstract' objects. Properly understood, and despite its unfortunate name, linguistic idealism is a resource materialists can use against all manner of spiritual entities, be they Platonic Forms or denizens of the 'third realm' of propositions, number and meaning postulated by logicians such as Frege.

Finally, I need to explain the connection between the present work and an earlier book In 1983 I published *Wittgenstein: A Social Theory of Knowledge* (Bloor 1983), which was a general survey of Wittgenstein's later philosophy, and a sustained attempt to

understand it as an exercise in the sociology of knowledge. This book differs in three ways. First, it is not a general survey, but deals exclusively with the single topic of rule following. In the former book I said little about rules, in this I talk of nothing else. I have resisted all temptation to expand the discussion. Inevitably there is a small overlap between the two books, but this is mainly the result of ensuring they can be read independently. Second, I previously took it for granted that when Wittgenstein referred to customs, conventions and institutions, he meant social customs, social conventions and social institutions. Unfortunately the climate is now such that it is necessary to argue the point.[1] Individualist presuppositions and individualist readings of Wittgenstein have become, not merely prominent, but dominant. Third, of the literature on Wittgenstein published since 1983 perhaps the bulk has been about rule following and the controversy between individualists and collectivists. I can now bring my earlier reading of Wittgenstein into contact with the current discussion. Much of it is extremely penetrating though, I have to confess, occasionally I find some of it perverse. I have not attempted a general survey, but have chosen to structure the discussion by confronting in detail one representative, book-length exposition of each side. Other work is discussed where necessary, but the collectivist case is (initially) represented by Saul Kripke's, *Wittgenstein on Rules and Private Languages* (1982), and the individualist counter-argument mainly by Colin McGinn's, *Wittgenstein on Meaning* (1984). Though it would be easy to find other vigorous statements of the rival positions to provide such a framework, these two effectively capture the main arguments currently under discussion, accurately reflecting both their strengths and weaknesses.[2]

Responding to the sheer quantity of all this work of commentary and analysis, J. J. C. Smart sounded a note of despair, 'I get the impression that practically everything that could sensibly be said on the topic, whether true or false, must have been said by someone, somewhere or other' (Smart 1992: 123). Smart is quite wrong. As we shall see, there are remarkable gaps and missed opportunities in the existing literature.

ACKNOWLEDGEMENTS

The stimulus for the present book came from an invitation to give two lectures on Wittgenstein at the Institute of Philosophy in Copenhagen. I must apologise to my hosts for the inordinate length of the reaction time between the stimulus of their invitation, in 1986, and the final response, in the form of this book. There were, alas, many intervening variables. Since that original engagement with the problem, I have had the opportunity to explore several of its many facets at seminars given in Cambridge, Edmonton, Florence, Jerusalem and Warsaw. Some of the material presented at those meetings has found its way into this book, some has been published elsewhere, and some has fallen by the wayside as a result of well aimed criticism from the other participants. I am particularly grateful to the members of the Science Studies Program at the University of California, San Diego, who in 1992 made possible a three month visiting professorship in the Department of Philosophy. This enabled me to offer a graduate course on Wittgenstein and to discuss some of the material presented here in a climate as invigorating intellectually as it was glorious meteorologically.

One special intellectual debt I must acknowledge is to a paper by my friend and former Edinburgh colleague, Barry Barnes. The paper is called, 'Social Life as Bootstrapped Induction' (Barnes 1983). I have made liberal use of it in Chapter 3 and would urge its importance on all philosophers who are interested in rule following. Several Edinburgh colleagues (past and present) read a complete draft of the book. I should particularly like to thank Barry Barnes, Celia Bloor, Martin Kusch, Peter Lewis, Donald Mackenzie and Steven Shapin for their valuable mixture of criticism and encouragement. It is also a pleasure, and an important duty, to thank the

two anonymous publisher's readers for their speedy and detailed comments. Their time and trouble certainly saved me from some confusions. I have done what I can to meet their criticisms though, on some fundamental issues, I have stubbornly refused to budge. I have also corresponded with Meredith Williams, Thomas Uebel and Nigel Pleasants. To do justice to the full range of issues raised by all these kind critics would, I fear, need another book. Last, but not least, I am indebted to Carole Tansley who typed a complete early draft and numerous redrafts of certain chapters.

ABBREVIATIONS

WITTGENSTEIN'S WORKS

BB	*The Blue and Brown Books,* 1969
CV	*Culture and Value,* 1980
OC	*On Certainty,* 1969
PI	*Philosophical Investigations,* 1967
RFM	*Remarks on the Foundations of Mathematics,* 1978
Z	*Zettel,* 1967

OTHER WORKS

AP	Thomas Reid, *Essays on the Active Powers of Man*, 1788
HA	Ludwig von Mises, *Human Action; A Treatise on Economics*, 1963
K	Saul Kripke, *Wittgenstein on Rules and Private Language*, 1982
M	Colin McGinn, *Wittgenstein on Meaning*, 1984
P	Karl Popper, *The Open Society and its Enemies*, 1969, vol.ii Hegel and Marx
T	David Hume, *Treatise of Human Nature*, 1739–40

REFERENCING CONVENTION

(i) Much of Wittgenstein's published writing is in the form of numbered paragraphs, thus PI:188 means *Philosophical Investigations* paragraph 188, and RFM II:6 is to be read as *Remarks on the Foundations of Mathematics*, part II paragraph 6. Where there is any danger of ambiguity, as in the short second part of the *Investigations*, a page number will be indicated, so PI II: p.224 means *Philosophical Investigations*, part II, page 224.

(ii): For works other than Wittgenstein's, one of the above letters, or pairs of letters, followed by a number indicates a page reference. For example, K:109 means Kripke's *Wittgenstein on Rules and Private Language*, page 109.

(iii): For full bibliographical details see the list of works cited in the Bibliography.

1

INTRODUCTION

Everyone is familiar with the sequence of even numbers: 2, 4, 6, 8, 10, 12, 14, and would know how to continue it. We use it when we are counting in twos, and no one has much difficulty with that. The sequence can be generated by following a simple rule: start with 2 as the first number in the sequence, and then add 2 to produce the next number of the sequence, and then add 2 to that for the next, and so on. The performance of following this rule is a trivial arithmetical achievement and yet, despite its commonplace character, isn't there also something strange and remarkable about it? For example, if we ask where the sequence ends, we have to say it doesn't. No one could reach the end because there is no end. If we are being solemn we might say the sequence is 'infinite'. So here is something apparently simple, following the rule 'add 2', and yet we immediately find ourselves rubbing shoulders with mysterious things like 'infinity'.

That isn't all. Think of the transition from one number of the sequence to the next: for example, when we take the step from 12 to 14, or from 14 to 16. If we are going to follow the rule we *have to* say 14 after 12; we *must* do it this way if we are to obey the rule. This is how we talk: we say 'have to' and 'must', but what do these words mean? What is it to '*have to*' say 14? If we '*must*' say 14 after 12, where does the 'must' come from? Does it mean we find it physically impossible to say anything else? Obviously not, because we could easily violate the rule and say 15 rather than 14, if we wanted to. So does 'must' mean we *want* to say 14? That doesn't sound right either: it surely isn't a case of wanting or not wanting. The point is, we have to say 14 because this is what the rule requires. But what does, 'the rule requires' mean? We know, or take ourselves to know, what it is for the government to require us to pay our taxes, or for parents to require their young children to be home in time for tea, but what is it

for a *rule* to require something of us? Perhaps it is the same kind of thing, but what kind is that? As Wittgenstein put it, 'You say you must; but cannot say what compels you' (RFM VI: 24).

We now have two mysteries: the mystery of infinity and the mystery of the compulsion of rules, or what Wittgenstein called the 'hardness of the logical *must*' (RFM I: 121). An interesting feature of this 'must' deserves special notice. If a rule of the kind we have just considered tells us we must do a certain thing, it is clear we are not being imposed on by a physical necessity which forces us to do it. Rather we must do the thing in question if we are to conduct ourselves rightly. We can do something else, but then we do something wrongly. So the necessity we are dealing with is like a moral necessity: it is to do with getting something right or wrong, and of behaving well or badly according to some standard embodied in the rule. We are in the realm of norms. In what follows I shall therefore talk of the 'normativity' of rules. When I use the word 'normativity' it is to be understood in reference to the considerations I have just outlined. The mystery of compulsion is thus the mystery of normativity.

If there is indeed something of a mystery about rules then it is a widespread one, because it will not just apply to number sequences: it will affect all comparable rules, and such rules are everywhere in our life. We obey or disobey rules of the road, rules of morality and decorum, rules of many different kinds of game, rules of grammar, rules of clubs and associations, as well as the rules governing numbers and numerical processes of the kind represented by our example. It is true that many of these rules may be less rigorous and exacting than the rules governing numbers, but they will share many of their essential features. In their various ways, and under various conditions, they will all 'require' something of us, and the requirement will typically apply to an open-ended and indeterminately large number of cases. If we want to understand rules in general we had better make sure we understand the clear-cut and simple cases – hence the example of the number sequence. Conversely, if we can get to the bottom of the mystery of the number sequence, we should be better placed to understand the nature of rules in general.

These lines of thought might make someone of a practical turn of mind impatient. It was admitted at the outset that following the rule 'add 2' was trivially easy, so why raise these issues and ask these questions? This refusal to be drawn into speculation about what I have called the mysterious aspect of rules is by no means confined to busy and practical people. It can also be found where it is least to

be expected, amongst philosophers whose business it is, one might think, to reflect on what the rest of us take for granted. In this context the impatience takes the form of saying the mysterious character of rule following isn't mysterious at all: it is just how things are. It must be accepted as a plain matter of fact. Our attitude should be one of 'relaxed common sense' (McDowell 1991: 169). Rules just do compel us, because if we are going to be consistent with the rule then we must do what it says on pain of violating the meaning of the rule – (all on the supposition, of course, that we are intending or meaning to follow it). In the case of the 2, 4, 6, 8 sequence, it will be said that once we fix the interpretation of 'two' and 'add' and 'sequence', and the other relevant words, then what we must do to follow the rule is completely determined. We are committed to saying 14 after 12. Both the compulsion and the endlessness of the rule cease to be mysterious, according to this argument, because they simply follow from what it is to mean something of this kind.[1]

Although he didn't embrace this view Wittgenstein characterised it accurately when he wrote:

> I feel that I have given the rule an interpretation before I have followed it; and that this interpretation is enough to *determine* what I have to do in order to follow it in the particular case (RFM VI: 30).

Let us adopt the terminology used in this passage and call this view, and the ideas associated with it, by the name of *meaning determinism*. This is the claim that the compelling and infinite character of rules derive from the property called 'meaning': i.e. the meaning of the rule itself and what is meant or intended by the rule follower. If someone were to give us the order 'start with two and count on in twos', then the meaning of the words in the order, and the meaning we attach to them, lays down what steps are to be taken. The order certainly doesn't mention, say, the step from 14 to 16, or from 1,000 to 1,002, but nevertheless its meaning is that we are to take precisely these steps. The meaning implies this in advance of our reaching the part of the sequence where we make this transition. Again Wittgenstein conveys the idea vividly:

> your idea was that that act of meaning the order had in its own way already traversed all those steps: that when you meant it your mind as it were flew ahead and took all the

> steps before you physically arrived at this or that one. Thus you were inclined to use such expressions as: 'The steps are *really* already taken, even before I take them in writing or orally or in thought' (PI: 188).

The idea of the steps having been 'already taken', of the right steps already existing in advance, is the nerve of what I have called 'meaning determinism'.[2]

Meaning determinism is an essential ingredient of a further position which will play a central role in the discussion to follow, namely individualism As I mentioned in the preface, this is one of the two fundamentally opposed approaches to rule following, the other being collectivism. Individualists do not deny that many rules are designed to regulate social interactions (e.g. the rules of etiquette). The point they insist on is that rule following as such doesn't necessarily or always involve interacting with other people. They express this by saying that rule following is not 'essentially social'. Some activities are essentially social, e.g. marching in step or following fashion. The rule for generating number sequences, however, says the individualist, is not of this kind. Obviously the rule itself doesn't refer to people but nor are other people involved in the following of it. We don't need other people to agree with us before we can be said to have got it right. For the individualist someone could, in principle, follow the sequence correctly even if everybody else disagreed or if nobody else were alive to reassure them or endorse their conclusions, or express their agreement. Even the fact that we are taught to add makes no difference to the individualist's claim. The need for teachers is said to be just a contingency, a non-essential fact that could in principle be otherwise. An individual could invent addition, or invent the sequence, rather than being taught it, and presumably at some point in the past something of this sort must have happened.

The justification for these individualist claims refers back to meaning determinism. It is based on the idea that rule following is made possible by our power to grasp the meaning of the concepts used in the rule. Once we have grasped them, then it is their meaning which guides or determines our behaviour. Grasping a concept is a purely individual achievement. It is an individual mental act or it is nothing. If we can do it at all we can do it for ourselves and (at least in principle) by ourselves. Teachers and helpers, though useful in practice, are providers of hints and

prompts rather than being logically necessary or involved by definition. Meaning determinism, then, purports to give us a picture of how the individual is provided with standards of right and wrong for their rule following activity: they are implicit in the very meanings of concepts. Without meaning determinism the individualist would have no account of normativity.[3]

Wittgenstein rejected meaning determinism and the various pictures and assumptions going along with it. He thought we could do better than this: we don't have to adopt the pretence that rule following is a simple, impenetrable matter of fact. We should let it arouse our empirical curiosity. He treated the facts of rule following, especially their more mysterious aspects, as things which can be opened up for investigation and which can be analysed in a revealing way. His theory was a form of collectivism and was extremely simple in its general outline. It may be summed up in the following two propositions: (1) rules are social institutions or social customs or social conventions; (2) to follow a rule is therefore to participate in an institution and to adopt or conform to a custom or convention. As he put it: 'To obey a rule, to make a report, to give an order, to play a game of chess, are *customs* (uses, institutions)' (PI: 199).

This is what a rule is, and all it is. It follows that explanatory categories dealing exclusively with our individual cognitive abilities will not be adequate to explain the phenomenon. More than individual psychology is needed. We will also need categories and theories drawn from sociology, even in the case of the 2, 4, 6, 8 sequence with which we began. The air of mystery surrounding rules, it will transpire, derives from our not having a clear apprehension of the social processes in which we ourselves participate. We are too close to recognise them for what they are, and therefore experience our institutions, customs and conventions in a distorted and transfigured form (cf. PI: 36, 196).

Can the notions of custom, convention and institution really tell the whole story? Isn't it obvious that a sociological, or conventionalist, or collectivist approach of this kind can only account for *part* of what goes on when we follow a rule? It may indeed be a matter of convention which rules we adopt, but surely once we have adopted them, what counts as correctly following the rule is fixed by the rule itself. It might be a matter of convention that the game we call 'chess' has the rules it does, but once they have been accepted it is their logical consequences which come into operation, determining

which moves conform to, or violate, the rules of the game. It is certainly a matter of convention that the word 'two' stands for the number two, and 'four' for the number four. It could have been the other way round. But once it has been agreed what the signs mean, then surely their logical implications are fixed. If this objection is correct, the 'must' and 'have to' of rules do not have their source in human conventions and institutions, but in consistency of meaning, and hence in the logical not the sociological properties of rules.

Clearly this is just a re-affirmation of the ideas already labelled 'meaning determinism'. If a sociological or conventionalist approach is to be sustained, the presuppositions of meaning determinism, and its associated individualism, must be met head on and replaced by a different conception of meaning. This is what Wittgenstein did: he developed an alternative account of meaning precisely for this purpose. For reasons which will become evident in the next chapter, the Wittgensteinian approach will be called *meaning finitism*.[4] To carry through the sociological programme care must be taken that the old assumptions about meaning are not smuggled in by the back door. It is all too easy to imagine a 'conventionalist' account of rules being accepted in a glib way without a clear sense of what is to be demanded of it. It is no use assenting to the proposition that a rule is a convention, if 'convention' is just another word for a rule, and rules are then given the standard, uncritical analysis of common sense – or its philosophical counterpart, the theory of meaning determinism. Sometimes Wittgenstein expressed his approach to rules by talking, not of custom and convention, but of 'norms' (e.g. RFM VII: 6I, VII: 67). I suspect that the trivialising tendency I have just described is not uncommon in connection with so-called 'norms'. Sociologists often analyse social structures by reference to a set of 'norms' where, for all practical purposes, a norm turns out to be a maxim or precept or verbal formula of some kind. These norms allegedly control behaviour by being 'internalised'. Once they have been internalised we are then taken to be under the control of the 'implications' of the norms. It will be clear that such an account is dangerously close to saying it is the meaning of the norm, the logical content of the maxim and its implications, that determines behaviour. Normative determinism, as it is called, is then no more than meaning determinism expressed in a sociological idiom.[5] Formulating an intellectually interesting sociological theory of rules, therefore, is no easy matter, given the dangers of lapsing back into an uncritical

meaning determinism, or accepting a mere variant of that doctrine in terminological disguise. We will see that Wittgenstein's meaning finitism avoids these traps. For this reason, Wittgenstein was more consistently sociological than many self-professed sociologists.

The idea that Wittgenstein developed a sociological account of rules is no novelty – and nor should it be given the arguments and formulations he used. When Wittgenstein's later philosophy was first published after his death in the 1950s it was read by supporters and critics alike as a social theory. The claim was taken to be that the compulsion of rules comes from the shared language games in which they played a part. Wittgenstein's idea, that the meaning of a sign came from its use, was taken to refer to a collective use and hence a collective body of users. The practices and customs alluded to by Wittgenstein were taken to be, self-evidently, social practices and social customs. When Ayer and Strawson criticised Wittgenstein they did so from an individualist standpoint, confident they were confronting an anti-individualist position (Ayer 1954; Strawson 1954). When Malcolm, Rhees and Winch defended Wittgenstein against these attacks they did so as collectivists defending a collectivist position (Malcolm 1954; Rhees 1954; Winch 1958).

Perhaps the most notable recent sociological reading, and one that has been at the centre of much controversy, is Saul Kripke's brilliant *Wittgenstein on Rules and Private Language* (1982). Although Kripke expresses reservations about the truth of Wittgenstein's conclusions, he has no doubts about the importance of his arguments, or their sociological character. Where will the present defence of collectivism, and the collectivist reading of Wittgenstein, stand in relation to Kripke's formulation? I am in sympathy with the main thrust of Kripke's argument, both as a positive account of rule following (despite his own reservations) and as an exposition of Wittgenstein's views. I shall seek to carry that argument forward. My analysis, however, will differ from Kripke's in five ways. First, Kripke does not fully bring to the surface Wittgenstein's finitist account of meaning. In particular, that account needs disentangling from certain themes Kripke chooses to emphasise, namely Wittgenstein's alleged shift from a correspondence theory of truth to a more pragmatic assertion condition approach. The basic ideas of finitism are all there in Kripke's account, but stand in need of greater systematisation. Second, his formulation of Wittgenstein's argument unfortunately contains a disconcerting and damaging ambiguity. To the best of my

knowledge this ambiguity has not previously been remarked upon, or if it has, it certainly has not attracted the attention it deserves in the literature. The effect of the ambiguity is to make the sociological approach in general, and Wittgenstein's version in particular, look more vulnerable to individualist criticism than they really are. Third, Kripke's exposition also contains a fault of the opposite tendency to the previous one: it transmits rather than repairs some of the genuine shortcomings in Wittgenstein's analysis of rules. As I shall show, there are ways in which that analysis needs strengthening. Fortunately this can be done in a manner that flows naturally from its sociological orientation. Fourth, I shall be taking up a hint dropped by Kripke in the very last footnote of his treatment of rules. Here he briefly compared Wittgenstein's argument to, of all things, Ludwig von Mises' attack on the economics of socialism. This striking and unusual comparison seems to have been passed over without comment in the literature. I propose to examine it and use it, in some measure against Kripke, to support my own claims about how to identify the real significance of Wittgenstein's argument. A fifth and final difference will be my exploration of a further historical parallel that casts light on the current controversy between individualists and collectivists. I shall compare currently opposed positions with those taken up by David Hume and Thomas Reid, respectively, in their debate about the origin and character of the obligation to keep promises. Kripke has already drawn a parallel between Hume's sceptical approach to causation and Wittgenstein's approach to rules. I shall not be addressing that particular comparison but argue for a different and, I believe, deeper connection between Wittgenstein's position and that of Book III of the *Treatise* (rather than, as with Kripke, Book I).

The immediate task is to set out clearly, and as simply as possible, Wittgenstein's positive account of the processes underlying what we call 'following a rule'. Different interpretations, opposing theories, and divergent readings can be dealt with when this primary duty has been discharged. In Chapters 2, 3 and 4 I will provide an exposition and development of Wittgenstein's position. It is here that the significant gaps in the secondary literature, mentioned in the Preface, will become apparent. We will then be in a position to assess the current confrontations and discussions as represented by Kripke and McGinn. Kripke will be discussed in Chapters 5 and 6, and McGinn in Chapters 7 and 8. The comparison with Hume and Reid will be introduced in Chapter 9

2

MEANING FINITISM

Wittgenstein said that if we want to understand rule following we should look at how we learn to follow rules, and how we might teach someone a rule. The circumstances of teaching and learning will give us all we need: 'Once you have described the procedure of this teaching and learning, you have said everything that can be said about acting correctly according to a rule' (RFM VII: 26).

There are, of course, descriptions and descriptions. A superficial description will surely not give us what we need. A revealing description will be one whose power to illuminate comes from an interesting theoretical perspective. Despite his frequently adopted (and frequently believed) posture of just 'describing' (cf. PI: 109, 124, 126) this is precisely what Wittgenstein had to offer. It is this theory we must recover and bring to the surface. With these qualifications the focus on teaching and learning is correct.

How, then, do we teach someone to follow a rule, e.g. the rule for producing the sequence of numbers 2, 4, 6, 8, 10, 12, etc.? Clearly we must first teach them to count. Then we must teach them to add and, in particular, to add 2. We will show them what we want them to do and encourage them to imitate, saying things like 'right!' and 'good' when they seem to be doing what we want, and 'no!' or 'wrong' to hold them back (PI: 208). We will even say, 'You must do this' and 'You can't do that' and shape the learner's behaviour by means of what Anscombe calls 'forcing modals' and 'stopping modals' (Anscombe 1978b; cf. RFM VII: 26). Initially these verbal formulations may be accompanied by physically guiding the learner's hands or by physically restraining them until, in the course of learning, purely verbal instructions produce the desired effect (RFM VI: 17, VII: 39).

So much is obvious, but Wittgenstein then draws attention to a

2

MEANING FINITISM

Wittgenstein said that if we want to understand rule following we should look at how we learn to follow rules, and how we might teach someone a rule. The circumstances of teaching and learning will give us all we need: 'Once you have described the procedure of this teaching and learning, you have said everything that can be said about acting correctly according to a rule' (RFM VII: 26).

There are, of course, descriptions and descriptions. A superficial description will surely not give us what we need. A revealing description will be one whose power to illuminate comes from an interesting theoretical perspective. Despite his frequently adopted (and frequently believed) posture of just 'describing' (cf. PI: 109, 124, 126) this is precisely what Wittgenstein had to offer. It is this theory we must recover and bring to the surface. With these qualifications the focus on teaching and learning is correct.

How, then, do we teach someone to follow a rule, e.g. the rule for producing the sequence of numbers 2, 4, 6, 8, 10, 12, etc.? Clearly we must first teach them to count. Then we must teach them to add and, in particular, to add 2. We will show them what we want them to do and encourage them to imitate, saying things like 'right!' and 'good' when they seem to be doing what we want, and 'no!' or 'wrong' to hold them back (PI: 208). We will even say, 'You must do this' and 'You can't do that' and shape the learner's behaviour by means of what Anscombe calls 'forcing modals' and 'stopping modals' (Anscombe 1978b; cf. RFM VII: 26). Initially these verbal formulations may be accompanied by physically guiding the learner's hands or by physically restraining them until, in the course of learning, purely verbal instructions produce the desired effect (RFM VI: 17, VII: 39).

So much is obvious, but Wittgenstein then draws attention to a

feature of the process which, though undeniable when pointed out, is easily overlooked. The number of illustrations and examples a teacher can offer a pupil must always be *finite*. By contrast, many of the rules that are being taught, including our 2, 4, 6, 8 sequence, are 'infinite'. Learning the rule and how to follow it isn't, therefore, like learning the alphabet, which is just a finite list. In these infinite cases the pupil has always got to go on beyond the given examples, or be deemed capable of going on, before they can be said to have learned what the teacher is trying to convey. This does not just apply to number sequences. Teaching someone the word 'red' is, in a sense, teaching them the rule for using the word. This too involves moving from a finite number of examples to an open-ended, indefinitely large range of future applications.

The point cannot be emphasised too strongly. In learning such rules there is always going to be the problem of taking the next step, of moving from previously known cases to new cases. The same applies to the day-to-day use of the rule. Here again we will be moving from known cases to new cases. The problem of taking the next step is always with us. It is this problem, the move from past to new instances of a concept, on which Wittgenstein wanted us to focus. We may express this by saying he wanted to make the step to the next instance *problematic*. He wasn't saying it will feel problematic to the rule follower. To an accomplished and well-trained rule follower it may feel smooth, automatic, and easy. But, regardless of how it feels, it is to be seen as problematic for any theory of rules. A proper account of rule following must focus the analyst's attention on that step in the process, i.e. the step to the next case.

This is not a new emphasis or a new insight. It is well established in the empiricist tradition. J. S. Mill, for example, canvassed the view that inference always carries us from 'particulars to particulars' (*Logic* 1848). He captured the standpoint of finitism well when he said: 'A name is not imposed at once and by previous purpose upon a *class* of objects, but is first applied to one thing, and then extended by a series of transitions to another' (*Logic* I: viii, 7).

Mill had in mind cases like 'red' rather than number sequences, but the point covers all concepts. Mill's idea of our extending the application of our terms 'by a series of transitions' is the one we need in order to think about rule following. Wittgenstein's own way of drawing attention to the problem of the next step was by pressing the question: how do I know what to do at *this* point? (RFM VI: 29). For example, confronting the picture of meaning

determinism, of the rule implicitly specifying what is to be done in advance, Wittgenstein posed the question: but how does that help me now? (i.e. in taking *this* step). A characteristic statement of the issue is this:

> *How do I know* that in working out the series +2 I must write '20004, 20006' and not '20004, 20008'? – (The question: 'How do I know that this colour is "red"?' is similar.) 'But you surely know for example that you must always write the *same* sequence of numbers in the units: 2, 4, 6, 8, 0, 2, 4, etc.' – Quite true: the problem must already appear in this sequence, and even in *this* one: 2, 2, 2, 2, etc. – For how do I know that I am to write '2' after the five hundredth '2'? i.e. that 'the same figure' in that place is '2'? And if I know it *in advance*, what use is this knowledge to me later on? I mean: how do I know what to do with this earlier knowledge when the step actually has to be taken? (RFM I: 3).

It might be objected that teaching a rule *doesn't* always reduce to showing the learner a finite number of examples. Sometimes we instruct learners verbally, and if the would-be rule followers understand our instructions they will be able to follow the rule by following the instructions. Someone might explain to us that the 'square' of a number n meant $n \times n$, and is written n^2. Instructed to produce the sequence n^2 for $n = 1, 2, 3, 4$, etc. we would then write *1, 4, 9, 16*, etc. We didn't do this on the basis of examples but through a verbal definition made up of words whose meanings are general and not confined to a finite number of cases. The reply to this objection will be obvious: we must first understand the words of the instruction or definition. If we ask how we learn these, we come back once again to examples (PI: 29). All learning therefore ultimately rests on finite numbers of examples, and this renders the problem of the move to the next step ineradicable.

To sustain the sense of the problematic character of the step-by-step movement in rule following Wittgenstein realised he had to root out alternative ways of glossing the process, i.e. ways that support meaning determinism. It is tempting to suppose that when a teacher is using examples to convey the meaning of a word, the teacher has something 'in mind', and the finite number of examples are just a fragmentary substitute for what is really meant. If only the pupil could look directly into the mind of the teacher then how simple life would be: they too would have access to the state of

understanding that is the source of the teacher's ability to follow the rule. Of course, they can't look into the teacher's mind but, the argument goes, they will only have established an understanding when they can reach beyond all the examples (RFM VII: 52). On this view, the source of understanding is qualitatively different in its nature from the examples themselves:

> Your idea, then, is that you know the application of the rule of the series quite apart from remembering actual applications to particular numbers. And you will perhaps say: 'Of course! For the series is infinite and the bit of it that I can have developed finite' (PI: 147).

If the teacher is thinking of the rule for counting in twos, and this is what he is trying to teach, then he has an idea of something going on for ever. The idea of the rule in the teacher's mind doesn't confine itself to a finite number of cases. It refers to, and captures, the infinite scope of the rule, but obviously he can't lay *that* before the pupil. So he contents himself with a poor second best – say, the first 10 or 20 members of the sequence, which he can lay out for inspection.

Wittgenstein conceded this picture may be 'entirely natural' but nevertheless rejected it (RFM VII: 52). For instance, he said teachers don't know more than they can say or make explicit. Addressing the teacher he said: 'You do not yourself understand any more of the rule than you can explain' (RFM VI: 23). And thinking of himself in the role of such a teacher, he said:

> And again I don't myself know any more about what I want from him, [i.e. the pupil, DB] than what the example itself shows. I can of course paraphrase the rule in all sorts of different forms, but that makes it more intelligible only for someone who can already follow these paraphrases (RFM VI: 19).

Finally:

> If you use a rule to give a description, you yourself do not know more than you say. I.e. you yourself do not foresee the application that you will make of the rule in a particular case. If you say 'and so on', you yourself do not know more than 'and so on' (RFM IV: 8).

This is a vital step in the argument, so we had better be clear about it. We are told that when we say 'and so on' we don't ourselves know more than 'and so on'. What does this mean? The point, of course, is that we cannot possibly conjure up in our mind, or bring into consciousness, all of the cases to which the rule is 'meant' to apply. The content of our consciousness in this regard must be finite. We might even say the content of the teacher's consciousness just is identical to the verbal formula that is given outer expression. But surely, we might object, the words 'and so on' convey some meaning and refer to something. Is Wittgenstein saying that, in responding to them, our mind is a blank and we have no sense of their gesturing towards the continuation of the number series? I don't think Wittgenstein was denying these things and telling us our mind is blank and we feel nothing on hearing the words 'and so on'. The issue is how these reactions are to be analysed, and in particular whether they can be explained without recourse to an 'idea in the mind', or a 'content' that transcends all the examples available for teaching and learning.

In the place of this picture, which is just another manifestation of meaning determinism, Wittgenstein substituted a more down-to-earth account, using biological ideas. He said we should think of ourselves as having instinctive responses to the examples used in teaching (e.g. OC: 359). 'I want to regard man here as an animal' (OC: 475). Given our biological make up, if we are shown three or four samples of a colour, and perhaps some samples of other colours to define the boundary, then for many purposes we can be sent off to fetch an object of a colour that matches it. We can learn the rule for 'red' in this way. Similarly we can be shown two objects, say two nuts, and on being told 'that's two', find it a perfectly adequate definition (PI: 28). This is possible because we instinctively generalise. Under certain conditions we spontaneously take ourselves to be encountering the same thing again which we 'recognise' in the sense of treating it as similar to previous instances of (what we take to be) this kind. We take something as red (again) or two (again). We don't need to have any 'criterion' or justification for deciding this is 'the *same*' again: we just react in this way (RFM VII: 40). The absence of doubt attending these reactions isn't a sign of haste or superficiality but rather, 'something that lies beyond being justified or unjustified . . . something animal' (OC: 358, 359).

When we are confronted with a finite set of examples we do not extract from them any general idea, rather, we instinctively pass on

to what strikes us as the next step or the next case. There is no rational basis for this, nor one that can ever be formulated in terms of propositions assented to by the rule follower. There are no assumptions or presuppositions or hypotheses or interpretations or 'meanings' that are attached to the examples. For Wittgenstein the bottom line of the explanation 'is not an ungrounded presupposition: it is an ungrounded way of acting' (OC: 110). We go on from our training in the way we do because we have a set of dispositions or tendencies that happen to be activated in this way by the examples used in training. It is not a kind of insight or 'seeing', i.e. seeing what the examples 'mean', that lies at the bottom of the language game, it is a way of acting (OC: 204). See Malcolm (1982).

As well as the biological instinct account of our spontaneous reactions and responses, Wittgenstein also resorted to mechanical analogies. He likened the rule follower to a piece of machinery (RFM IV: 20). This is apt for explaining how a rule follower can go on from a finite number of examples, because mechanical systems have properties such as momentum and inertia. They can also contain parts which rotate or cycle around and repeat themselves. Once they are set in motion they continue (for a while) to go further along the same path, or to repeat again what they did in the past. This makes them obvious models for the repetitive processes that often attend rule following, especially of the number sequence variety. A person taught to produce such a sequence will often do so automatically, grinding out the numbers in a machine-like way. Being given a rule sets the machinery in motion. Wittgenstein even used the metaphor of rule users 'unwinding' (RFM III: 69) – that is, giving their automatic tendencies free reign, handing over to the tendencies they find inside themselves (RFM VII: 4).

We can now see why Wittgenstein posed his question about the utility of saying the correct steps of a rule are somehow laid out in advance. For him this supposition was unnecessary. If we respond automatically and instinctively, we don't need to know in advance.

> But do you mean to say that the expression '+2' leaves you in doubt what you are to do e.g. after 2004?' – No; I answer '2006' without hesitation. But just for that reason it is superfluous to suppose that this was determined earlier on. My having no doubt in face of the question does *not* mean that it has been answered in advance (RFM I: 3).

The point will bear emphasis. The idea that the correctness of the answer 2006, 'was determined earlier on' is rejected. It is superfluous. So it was not determined earlier on. If it is determined at all, it is determined at the moment the step is made, i.e. by the contingencies operative at that time. This is the rejection of meaning determinism in its most explicit form.[1]

Biological and mechanical models, then, can help explain why our response to a finite number of examples might contain within itself an innate tendency to go beyond them. There is one further and highly important aspect of the step from past cases to new cases yet to be dealt with. So far the account has said nothing about whether the move to the next putative instance of a rule or concept is right or wrong. Saying pupils extend their training in ways that seem natural to them, or in ways that come unhesitatingly, does not yet address, let alone solve, the mystery of normativity. We cannot say whatever is instinctive is going to be right. That would reduce norms to the merest subjectivity: 'Then according to you everybody could continue the series as he likes . . . !' (RFM I: 116). Suppose a number of would-be rule followers were trained to count in twos, and when they followed their instincts they found they diverged from one another. We want to say, and the rule followers themselves will want to say, that somebody was making a 'mistake' but, 'what is the criterion by which this is a *mistake*?' (PI: 51). As yet there is no provision in the model for this. In order to give normative notions some purchase Wittgenstein told us to reject any identification of standards of right and wrong with subjective tendencies or feelings. For the pure follower of instinct, or for those who would trust their automatic reactions, whatever course of action they feel welling up within them would be 'right' by definition. For Wittgenstein this spelled the total destruction of normative notions. 'One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about "right"' (PI: 258).

If the notions of right and wrong cannot be explained by the mechanisms within each individual leading them spontaneously from case to case, what can explain them? Wittgenstein's answer was that these normative ideas are bounded by 'the role of thinking and inferring in our life' (RFM I: 116). He meant that a tacit *consensus* of action determines what is counted as right and wrong. Making a step in following a rule counts as a 'right' step, i.e. a genuine and successful piece of rule following, if it is aligned with the steps

everyone else, or nearly everyone else, takes. Discussing the teaching of a rule generated sequence Wittgenstein said, 'let us suppose that after some efforts on the teacher's part he continues the series correctly, that is, as we do it' (PI: 145). Then, he said, we can say he has mastered the system. The operative words here are those identifying 'correctly' with 'as we do it'. Generalising from the number sequence to the rules of language, the same consensual theme is emphasised. Rules of language are based on regularity, on agreement in action:

> Here it is of the greatest importance that all or the enormous majority of us agree in certain things. I can, e.g., be quite sure that the colour of this object will be called 'green' by far the most of the human beings who see it (RFM VI: 39, cf. VI: 30, V11: 40).

To make a 'wrong' move is ultimately to make a move that leads the individual along a divergent path. To be wrong is to be a deviant, though this isn't to say the word 'wrong' *means* 'deviant'. It is necessary to introduce the qualification 'ultimately', and to say that *ultimately* to be wrong is to be deviant, because often such judgements are mediated by appeal to notions of authority and expertise. It is easy to see, however, that these are only complications in the story, not things challenging its basic principles. They merely raise the question of the grounds of authority and the identification of expertise, and this brings us back to some form of consensus. The normative properties of rules, then, do not derive from the instinctive sources of individual activity, but from the alignment of these different sources in the majority of cases, or what Wittgenstein called 'the agreement of ratifications' (RFM VII: 9).

Wittgenstein sometimes expressed himself by saying that consensus is a precondition of rule following activities, e.g. of arithmetical calculation: 'This consensus belongs to the essence of *calculation*, so much is certain. I.e.: this consensus is part of the phenomenon of our calculating' (RFM III: 67).

Because consensus is integral to following mathematical rules we can use these rules to make predictions or prophecies about how rule followers are likely to behave. If I ask you to multiply 25 by 25 I could predict you will say 625. Calculating, said Wittgenstein, is founded on such predictions (RFM III: 66); but obviously that doesn't mean that the point of calculating, or the concern of the typical calculator, is to make such prophecies. The interesting result

is not what someone is going to say, but what they ought to say, and that applies in our own case as well (RFM III: 69):

> The prophecy does *not* run, that a man will get *this* result when he follows this rule . . . – but that he will get this result, when we *say* that he is following the rule (RFM III: 66).

Our concern with the consensual aspects of rules is normative and evaluative. If our prediction or prophecy comes out wrong – we say he will write down 625, and he writes down 652 – we do not call the 'theory' on which the prediction is based into question. Rather we put the competence of the rule follower in doubt (cf. RFM VI: 5). The involvement of consensus at some point in the story was predictable from the start: it was implicit in the focus on teaching and learning. Teachers are the agents of the social collective who are charged with the task of transmitting its culture to new members. Wittgenstein was quite explicit and matter-of-fact in his handling of this theme. In connection with the rules of logic, specifying what is to count as an acceptable inference, he said the rules do not compel like rails compelling a locomotive: 'Nevertheless the laws of inference can be said to compel us; in the same sense, that is to say, as other laws in human society' (RFM I: 116).

He considered a 'clerk' given the military sounding task of going through a list of people's heights and assigning them, according to the regulations, to certain 'sections': 'The clerk . . . *must* do it like that; he would be punished if he inferred differently. If you draw different conclusions you do indeed get into conflict, e.g. with society; and also with other practical consequences' (RFM I: 116).

The argument so far may be summarised like this: in following a rule we move automatically from case to case, guided by our instinctive (but socially educated) sense of 'sameness'. Such a sense does not itself suffice to create a standard of right and wrong. It is necessary to introduce a sociological element into the account to explain normativity. Normative standards come from the consensus generated by a number of interacting rule followers, and it is maintained by collectively monitoring, controlling and sanctioning their individual tendencies. Consensus makes norms objective, that is, a source of external and impersonal constraint on the individual. It gives substance to the distinction between rule followers thinking they have got it right, and their having really got it right. Clearly, much more needs to be said about the way the appeal to rules, and references to rules, enters into the process of maintaining consensus.

For example, deviation is sanctioned in the name of the rule, and by reference to the rule – we must not think of the mechanisms for maintaining consensus as one thing and the rule as another. The idea of the rule enters into the understanding and apprehension and maintenance of the consensus itself. This is an important theme and will be explored in the next chapter. Before developing Wittgenstein's theory along these lines, however, I want to describe his criticisms of an important rival understanding of the process of rule following.

One idea that often features in accounts of rule following is 'interpretation'. We have already encountered it in the sketch of meaning determinism: a certain interpretation is put on the rule, and this interpretation is supposed to determine the response. The interpretation compels or guides the rule follower. What is an interpretation? An interpretation of a poem states what the poem states, but in different words. When an interpreter translates a German sentence, we are given another sentence, say, in English. One set of signs is replaced by another set of signs, under the constraint that the meaning of the second set is the same as, or very close to, the meaning of the first. So interpretation is not a process that generates meaning: it is a transformation that takes the notion of meaning for granted. The source of meaning cannot be interpretation itself (PI: 198, 201; Z: 234–5). This immediately creates a problem if we say we follow a rule by interpreting it, or if we say every act of rule following involves our interpreting it. While it will sometimes be true that we follow a rule by interpreting it, e.g. by restating it in a language we understand, or breaking down a procedure into a set of simpler more explicit sub-procedures, this cannot always be true. It cannot always be true because the process of interpretation itself will have the character of an act of rule following. After the interpretation we will still be left with a set of symbols to which we must respond, i.e. upon which we must act. Wittgenstein concluded:

> What this shows is that there is a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call 'obeying the rule' and 'going against it' in actual cases (PI: 201).

The original German formulation of the final words of this passage spoke of the application of the rule going '*von Fall zu Fall*', from case to case. This literal rendering perhaps captures more clearly

than the standard translation 'in actual cases' the finitist picture being advanced here.

Disastrous consequences follow from insisting that an act of interpretation always mediates a step in rule following. If to understand a rule I must interpret it, then I must interpret the interpretation (cf. Z: 229–30). If meaning is not something independent of interpretation, and a standard against which it is measured, then there is nothing to constrain the interpretation. These consequences are set out by Wittgenstein in the passage below:

> 'But am I not compelled, then, to go the way I do in a chain of inferences?' – Compelled? After all I can presumably go as I choose! – 'But if you want to remain in accord with the rules you *must* go this way.' – Not at all, I call *this* 'accord'. – 'Then you have changed the meaning of the word "accord", or the meaning of the rule.' – No; – who says what 'change' and 'remaining the same' mean here?
>
> However many rules you give me – I give a rule which justifies *my* employment of your rules (RFM I: 113).

We must be careful to keep track of whose side Wittgenstein is on in this exchange. Is he asserting or denying that anything can be derived from anything according to some rule – 'or even according to any rule, with a suitable interpretation' (RFM I: 7)? The answer is, he accepts that *whatever* we do can be interpreted as following a given rule, and that '*anything* can be somehow justified' (RFM VI: 38–39). This result confronts us with a choice: either there is no determinate phenomenon of the kind we call rule following, or interpretation plays no ultimate role in it. Wittgenstein opts for the latter conclusion. The ultimate non-interpretive step is something entirely non-intuitive, entirely non-self-intimating and non-self-justifying. The ultimate step, of course, is underpinned by the operation of those instinctive and mechanical response tendencies already introduced. The process has a causal, psychological terminus, not a logical terminus (Z: 231; PI: 140, 220). Thus: 'When I obey a rule, I do not choose. I obey the rule *blindly*' (PI: 219).

According to meaning finitism, we create meaning as we move from case to case. We *could* take our concepts or rules anywhere, in any direction, and count anything as a new member of an old class, or of the same kind as some existing finite set of past cases. We are not prevented by 'logic' or by 'meanings' from doing this, if by these words we have in mind something other than the down-to-earth

contingencies surrounding each particular act of concept application. (Some interpretive gloss can always be provided to render the step formally consistent.) The real sources of constraint preventing our going anywhere and everywhere, as we move from case to case, are the local circumstances impinging upon us: our instincts, our biological nature, our sense experience, our interactions with other people, our immediate purposes, our training, our anticipation of and response to sanctions, and so on through the gamut of causes, starting with the psychological and ending with the sociological. That is the message of Wittgenstein's meaning finitism.[2]

Referring to mathematical rules of the kind used to find approximate values of irrational numbers, such as the square root of 2, Wittgenstein said: 'I want to say: it *looks* as if a ground for the decision were already there; and it has yet to be invented' (RFM V: 9). He went on:

> And what is in question here is of course not merely the case of the expansion of a real number, or in general the production of mathematical signs, but every analogous process, whether it is a game, a dance, etc., etc. (RFM V: 9).

In other words, the finitist picture of rule following isn't confined to special cases, it is meant to apply everywhere, including those places such as mathematics where it would least be expected.

The most obvious feature of this picture – something so obvious it is in danger of being overlooked – is its wholly down-to-earth character. It is this-worldly, concrete and causal: in a word, 'naturalistic'. Wittgenstein insisted he was talking about language, even mathematical language, as a 'spatial and temporal phenomenon', not some abstract 'phantasm' outside space and time (PI: 108). One potent source of such phantasms is meaning determinism. What is the attraction of these non-finitist theories in which the mind allegedly grasps the content of ideas which lay out in advance the right and wrong applications of the concept? Where do we get the idea that a rule has some manner of independent existence, and that following a rule is like tracing out a line that is already there? As Wittgenstein put it, with regard to the rules that govern number sequences: 'Whence comes the idea that the beginning of a series is a visible section of rails invisibly laid to infinity?' (PI: 218). This isn't just a naive picture people grow out of as they become more intellectually sophisticated: it has also gripped philosophers. Wittgenstein said Frege was of this view:

> Thus Frege somewhere says that the straight line which connects any two points is really already there before we draw it; and it is the same when we say that the transitions, say in the series +2, have really already been made before we make them orally or in writing – as it were tracing them (RFM I: 21).

Wittgenstein's hypothesis was that we are prone to accept this picture because of the strong drill and rigorous training necessary if we are to master rules, especially mathematical ones. We feel there is an inexorability about the progression of rules such as the number sequence, because of the inexorability with which we are socialised into the practice. The emphatic but mysterious '*must*' corresponds to our collective attitude towards the rule (RFM VII: 67). The process of training is so rigorous because of its centrality and importance in our lives: 'And that is why we learn to count as we do: with endless practice, with merciless exactitude; that is why it is inexorably insisted that we shall all say "two" after "one", "three" after "two" and so on' (RFM I: 4).

In reality it is not some abstraction that is inexorable, it is '*we* that are inexorable' (RFM I: 118). We experience the rule itself as having the character that should rightly be attributed to our own collective behaviour. None of our talk about rules existing in advance of our following them should be taken literally. The things we feel inclined to say about rules, or the pictures that come into our heads when we reflect on them, are not genuine insights into rules or data about them. They are just the psychological side-effects of social causes and, as such, they should be seen merely as material for explanation or, as Wittgenstein put it, things for 'philosophical *treatment*' (PI: 254). They only make sense if they are 'understood symbolically': 'I should have said: *This is how it strikes me*' (PI: 219).

And he added: 'My symbolical expression was really a mythological description of the use of a rule' (PI: 221).

Thus the mysterious aspects of rules are dismissed as myths, but myths arising from our failure to bring our experience of social processes and pressures into focus. The claim is that when we find our rule following to be smooth and mechanical we project that feeling onto the rule itself: 'The rule can only seem to me to produce all its consequences in advance if I draw them as a *matter of course*' (PI: 238).

Admittedly, other experiences seem to support the idea of the independent reality of rules, and suggest it isn't a myth. We sometimes get flashes of insight, e.g. into how a number series is to continue, or how a tune goes that we have been trying to remember. We might want to say of this experience, 'It's a particular feeling, as if it *were there*', but, Wittgenstein asked, do we have a model for this? He said that this sort of talk is nothing but a turn of phrase that suggests itself to us, and should not be taken literally: 'You have no model of this superlative fact, but you are seduced into using a *super-expression*' (RFM I: 124; cf. also I: 123).

For Wittgenstein, rules and meanings considered in themselves do not possess any agency: all agency and action associated with them derives from their human users and creators. And yet we constantly speak as if we are compelled by some reality outside ourselves. This is not, however, pure error and illusion, for each of us individually *is* compelled by something outside, namely by other people around us in society. It is society that is external to us and the true source of our sense that rules exist as an independent reality set over against the individual rule follower. So there is a reality answering to these mysterious, myth-ridden feelings, but nothing that lies beyond the social collectivity and its constituent parts. We are only compelled by rules in so far as we, collectively, compel one another.

So far I have only discussed one of the two mysterious aspects of rules I identified in the first chapter, namely the mystery of normativity. What about the mystery of infinity? How can a rule generate, say, an infinite sequence of numbers? How is meaning finitism to be reconciled with our grasp of 'infinity'? When doing arithmetic we seem to have no difficulty handling propositions like, 'there is no greatest cardinal number'; nor do we hesitate before declaring that the sequence 2, 4, 6, 8, 10, etc. has no end. It seems we *do* grasp the fact some things go on for ever, getting bigger and bigger without end. This seems to confront us with a difficult choice. On the one hand we feel inclined to say the finite mind surely cannot grasp the infinite (RFM V: 6). On the other hand, we *do* seem to grasp the infinite. Could it be that our minds actually partake in some way in the infinities they can grasp, so our capacities are not really finite after all? Meaning finitists reject this option. Wittgenstein sustains the finitist position by treating 'infinity' as something negative rather than positive. Infinity is a mere lack, namely the lack of any specified limit or end. Things are left open by virtue of their *not* possessing certain properties rather than by their being heavily,

indeed 'infinitely', weighed down: 'To say that a technique is unlimited does *not* mean that it goes on without ever stopping – that it increases immeasurably; but that it lacks the institution of the end, that it is not finished off' (RFM II: 45).

For example, we might say the playing field for a certain game is 'unlimited', and mean merely that 'the rules of the game do not prescribe any boundaries – say by means of a line' (RFM II: 45). The situation might be likened to having *permission* to do something, with nothing being said about stopping. Such a permission isn't thereby freighted with overwhelming positive significance: it is simply silent about the end, it has nothing to say on the matter (RFM II: 26–27; cf. RFM V: 14, V: 19, V: 36).

If we examine the rule for generating the sequence of even numbers, or any other mathematical system or 'calculus', all we find on the page are a few words or a few signs. Looking at the words or signs (even the sign 'and so on', or the dots at the end of a finite sequence indicating its continuation), we see something perfectly concrete, delimited and (finitely) graspable. We don't, we might say, discover anything on the page (mysteriously) 'infinite'. We might express this by saying, 'But when one examines the calculus there is nothing infinite there'. Wittgenstein said this way of talking is clumsy; but clumsy or not, he thought it was on the right track. It is, he said, a way of asking: 'is it really necessary here to conjure up the picture of the infinite (of the enormously big)?' (RFM II: 59). His answer was that it isn't necessary. To react with disappointment at *not* finding anything infinite in the calculus or the rule or the system of signs on the page is, for Wittgenstein, just amusing. The correct response is to ask, 'what is the everyday employment of the word "infinite", which gives it its meaning for us?' (RFM II: 60). Its everyday use comes from our learning to gloss techniques that lack the institution of an end by calling them 'infinite'. In this way the mystery disappears, and the use of the word 'infinite' is seen to be consistent with meaning finitism. The important condition to be satisfied, and the crucial insight, is that the word 'infinite' must get its meaning *from* our finite, rule-following activities. Infinity isn't something having an independent meaning we (somehow) intuit and then bring *to* these activities as a specification they must fulfil (RFM II: 58). The (finitist) meaning of the word 'infinity' comes from its use, not its use from its (non-finitistically given) 'meaning'.[3]

One of the most basic concepts used in philosophical semantics is that of the 'extension' of a term. The extension of 'swan' is the

class of all things, past present and future, correctly called swans. The extension of 'water' consists of everything that might have been, or might be, truly classed as water. It is usually noted that the extension of a term is not to be equated with its meaning, because two terms with the same extension might have different meanings. Nevertheless the meaning of a term is said to fix the extension. Once the meaning is given the class of things within its extension has, allegedly, thereby been predetermined.[4]

For a finitist there is no such thing as the 'extension' of a term or concept, or, if the word 'extension' is used, it radically changes its significance. Thus, Wittgenstein said we can use the word 'number' so that 'the extension of the concept is *not* closed by a frontier' (PI: 68). Without (closed) extensions of the traditional kind, however, we no longer have propositions with a determinate content, at least, not in the fashion usually pictured. There is no class of things existing in advance of the application of a label. Here and now, there is no determinate class of things which will, or could, truly be called swans. The content of that class depends on decisions which have yet to be taken, and so does not yet exist. Particular things, or individual objects, exist in advance, but not classes of things. The claim is that 'extensions', as philosophers have characterised them, don't exist. They are simply fictions generated by a philosophical theory. When philosophers think of the 'extension' of a term they think of an envelope surrounding a set of objects. They imagine a line drawn round them delimiting a definite area, rather like the Venn diagrams used in logic and set theory. Thus, 'Frege compares a concept to an area and says that an area with vague boundaries cannot be called an area at all' (PI: 71). Wittgenstein returned to the theme using a slightly different metaphor when he said:

> The sense of a sentence – one would like to say – may, of course, leave this or that open, but the sentence must nevertheless have *a* definite sense. An indefinite sense – that would really not be a sense *at all*. – This is like: An indefinite boundary is not really a boundary at all. Here one thinks perhaps: if I say 'I have locked the man up fast in the room – there is only one door left open' – then I simply haven't locked him in at all; his being locked in is a sham. One would be inclined to say here: 'You haven't done anything at all'. An enclosure with a hole in it is as good as *none*. – But is that true? (PI: 99)

Following Wittgenstein, the finitist will reply that it isn't true: an enclosure with a hole in it can be perfectly serviceable for some purposes – or, to put the point in another way, when talking about concepts and meanings the metaphor of enclosure is not appropriate.[5]

From a finitist perspective it makes sense to say our beliefs are never determinate or fully specifiable, at least, not in the way many philosophical theories have previously assumed. Nor are our intentions ever fully specified with respect to the circumstances that might be counted as fulfilling or frustrating them. To state the matter in the most uncompromising way possible: we can never know or state exactly what we believe, or what we intend. Of course, this sounds wrong because it draws a contrast with an unrealisable philosophical ideal, but the important point is not phraseology. What is important is the recognition that we do not, and cannot, plan in advance for all contingencies, and yet in practice we never feel this fact as a lack. In this respect Wittgenstein surely intended his famous discussion of the concept 'game' to be fully representative.

> For how is the concept of a game bounded? What still counts as a game and what no longer does? Can you give the boundary? No. You can *draw* one; for none has so far been drawn. (But that never troubled you before when you used the word 'game'.) (PI: 68)

Wittgenstein is not saying that the word 'game' is defective, or that we can't construct such expressions as 'all games' or 'the class of games'. But he is saying that when we invoke them an irreparable indeterminacy hovers around their employment. Likewise for all our everyday and scientific employment of class and kind terms. We do and can talk about kinds of things. Such talk does not refute finitism, but it must be analysed finitistically. And the same applies to the boundary that can be (artificially) drawn around games or the class of numbers. We could say: these, and only these, are truly games or numbers, and lay down a set of specifications. But the concepts employed in the specifications would themselves be subject to an inherent indeterminacy. This is what links Wittgenstein's discussion of the vagueness and open-texture of concepts such as 'game' or 'number' with his finitism. Cases of vagueness which we once lived with, and of which we may have been wholly unaware, can act as reminders that an enclosure with a hole in it is not without its uses. The thesis that all enclosures always have 'holes' in

them – by traditional philosophical standards – then looks less like a non-starter. There is a developing thread of argument running through the *Investigations* which gradually introduces these themes and carries us from the easily accepted phenomenon of open-texture to the radical thesis of finitism (cf. Bloor 1996). Wittgenstein's aim was to make us aware that we simply do not address the question of what we should say, or how we should apply or withhold a class or kind word, if objects behaved in ways that we did not expect or anticipate (cf. PI: 79–80). These are, in the most profound and important sense, open questions. What is more, they are just as much open questions for the social collective as for the individual. Finitism doesn't imply that if you examine individual thoughts you will find meaning is indeterminate, but if you bring in the community this indeterminacy is removed or corrected. It can never be removed. Consensus may furnish us with norms, but it does not overcome finitism. Nothing can overcome finitism.

3

RULES AS INSTITUTIONS

The three dimensions of Wittgenstein's theory of rule following have now been identified. They are (1) its biological or psychological aspect, dealing with our instinctive and automatic responses, (2) its sociological or collectivist aspect which concerns the shaping and sanctioning of our innate tendencies and their organisation into customs, conventions and institutions, and (3) the background of meaning finitism against which the entire process is set. All three must be kept in mind when we respond to Wittgenstein's slogan: 'A game, a language, a rule is an institution' (RFM VI: 32). Our understanding of the claim that rules are institutions will, however, also depend on how deep or how shallow an idea we have of institutions. We need more than a generalised awareness of the importance of social processes: we need a specific understanding of what is meant by the word 'institution'. The same applies to Wittgenstein's references to conventions and customs. As I indicated in the previous chapter, so far the statement of the sociological aspect of Wittgenstein's theory has been relatively undemanding. It is now time to probe more deeply. We must find an illuminating way to characterise the central, sociological concepts on which the theory depends.

Here we must brace ourselves for a shock. Wittgenstein at no point explained or defined the words 'custom', 'convention' or 'institution'. He treated them as he treated the other basic terms of his analysis (such as 'game') as well-understood words of ordinary language. If we examine passages where the notion of 'custom' is invoked, all we find by way of amplification is a reference to regularities in behaviour and the repetitive character of customs. Consider what is said about following a sign-post. For Wittgenstein a sign-post is an example of a rule, and following a sign-post is a case of following a rule. One way (though not the only way) to link

the sign-post example with the finitist account of rules is as follows: think of the sign-post as like the initial instruction, or the initial statement of the rule, in a number sequence generation problem. The sign-post says 'this way'; the instruction says '+2'. Just as we set off generating numbers, so we look at the sign-post and then journey on, leaving the sign-post behind, but still (as we say) 'following' it. How do we do this? Well, said Wittgenstein, perhaps we have been trained to act in certain ways to such signs, and now we do so act (PI: 198). He denied that this gives us a 'merely' causal account or that it by-passes the essential nexus of 'following' or 'going by' the rule: 'On the contrary; I have further indicated that a person goes by a sign-post only in so far as there exists a regular use of a sign-post, a custom' (PI: 198).

It would not be proper, he argued, to say somebody followed a sign-post just once in the history of mankind. We can say they once walked parallel to a board, but not that this was an act of 'following' (RFM VI: 43). So 'custom' implies (at least) a regular use and precludes single, isolated episodes by single individuals (cf. also RFM III: 67, VI: 21).

These conclusions are plausible and predictable. They are also disturbingly minimal, hardly adding depth to our understanding of the theory or taking us beyond the formulations of the previous chapter. If we are to deepen the analysis it is necessary to find some new resources and add some new ideas. Thus, to understand the notion of 'convention' we could appeal to Hume's definition, and I shall be doing that in a later chapter.[1] Perhaps the most pressing need is to develop, at least in outline, an account of the nature of institutions. Since Wittgenstein uses the idea, but doesn't offer such an account, we are forced to leave his text behind for a moment, and take the next step without its immediate guidance. In these circumstances one might have expected the secondary literature to be full of candidate accounts designed to make good the deficit, but this is not the case. Positive ideas on the subject have been conspicuous by their absence.[2] Fortunately there are resources available to fill this gap. I shall be making use of the ideas of two writers, one a philosopher (Anscombe 1969, 1976, 1978a, 1978b) the other a sociologist (Barnes 1983). Their work can be used to provide us with a simple model of an institution. After taking this wider look round, and seeing what they have to tell us, the intention is to return to Wittgenstein's argument. We should then be in a better position to respond to it in an informed way. Let us, then, address the question directly: what is an institution?

We may start by identifying some uncontroversial cases of social institutions. Clearly, money is an institution, marriage is an institution, and property is an institution. Reflecting on these cases might reveal to us general characteristics that can then be discerned in the case of rule following. Think of a simple economic system in which 'coins', in the form of metal discs, are used to facilitate exchange. We cannot say a coin simply *is* an appropriately shaped and stamped metal disc. Such objects could exist without being coins: they might be ornaments. Coins have shape and physical substance, but being a coin isn't a matter of shape or physical substance. It is a matter of how the thing is used. We discover the character of a coin, as a coin, not by examining *it* – by studying its geometry or physics or chemistry – but by seeing how people relate to it. The important thing is how people regard it and employ it as a medium when interacting with one another. We must attend, not to the thing itself, the thing we call a 'coin', but to the people who call that thing a 'coin'. I shall express this by saying metal discs *are* coins because they are *called* 'coins'. My references to things being *called* 'coins' is shorthand for the entire repertoire of behaviour associated with their being thought of, or regarded as, or treated as, coins (cf. RFM I: 129). On this usage, speaking of a thing as a coin isn't meant to refer to a purely verbal act, but to the whole pattern of behaviour into which such explicit verbalisations are woven. Bearing in mind the compression involved, the important point is that the group practice of calling a certain type of object a coin, makes that object into a coin.

Looked at in one way this is unremarkable; looked at in another way it is a striking exercise in pure creativity. A group of people have brought something into existence simply by thinking about it and talking about it. They have created coins in a way that looks almost magical. It is not magic but it is certainly a case in which an 'idealist' theory seems appropriate and correct. Idealists assert that reality has a spiritual or mental essence, and here we have money being constituted by being referred to as money, and having no existence outside of that practice:

> But what is produced is not an independently describable effect, as it were magically brought about by signs. It does not exist at all except for the signs. It is as if words produced it by signalling it.
>
> (Anscombe 1978a, reprinted 1981: 138)

The tree outside may still be there when I stop thinking about it, but if we collectively abandoned all reference to money it would vanish. (If it were said in objection that money would *not* vanish in these circumstances, because the actual coins would still be there, we can see that the social identity of a thing is being confused with its non-social, material nature. 'Actual coins' are not to be equated with the metal discs which certainly would survive the hypothetical abandonment of monetary practices. The discs only become coins, and hence money, in the context of the institution of money. It is not an exaggeration, then, to insist that if we abandoned references to money, money would indeed vanish.) We can see why Anscombe introduced the label of 'Linguistic Idealism' to describe the analysis appropriate to cases of this kind (cf. Bloor 1996).

Coins must be rather different kinds of things from – let us say – trees. Calling something a 'tree' might make the word 'tree' the name of that kind of thing, but it doesn't bring the object so designated into existence. Our experience of the world assures us that individual trees tend to be quite impervious to what is said about them. The kind, class or concept may be a human invention, but the objects falling under the concept, or belonging to the kind or class, are not human inventions. Instances of what may be called 'natural kinds' (e.g. trees, pebbles, cats, waves, molecules, etc.) are, in this respect, different from instances of 'social kinds' (e.g. coins, monarchs, masters, slaves, etc.) precisely because they have an existence independent of our regard. In the case of instances of natural kinds our thinking and talking and acting must be 'matched' in some pragmatic way against an independent, non-verbal reality. In the case of social kinds such as coins, however, our talk isn't so matched. Calling something a coin is correct because it is the practice to call it a coin. Although 'coin' doesn't mean 'called a coin by others', ultimately, it is correct to call something a coin because others call it a coin. If there is a reality that matches or corresponds to any given episode of such talk it can only be the totality made up of this and all the other similar episodes. Talk of coins, taken collectively, is not about a reality that is independent of such talk. It is, in a sense, just talk about talk. This is not how it appears to those who are immersed in the institution, for them the talk is about money, but viewed, as we are viewing it, from the outside this characterisation is correct as a description of the overall structure.

The same considerations apply to the example of property. Jones is the proud owner of a house. The house is his property, but 'being

owned by Jones', though truly predicated of the house, is unlike some of the other things predicated of it (like having a tiled roof and square windows). The latter are things found in or on the house.[3] We don't want to conclude that being owned by Jones is an immaterial, invisible, spiritual ingredient of the house, so what can we say? We can follow the model of the coin and say that, in the last analysis, being owned by Jones comes down to the fact that enough people are disposed to agree it belongs to Jones. It is true to call it 'owned by Jones' because others also call it (or if asked, would call it) 'owned by Jones'. What, however, is this consensus a consensus *about*? Its substance surely reduces to accepted patterns of behaviour oriented to Jones and the house, but if we try to state the 'meaning' of that behaviour or articulate the idea that informs it, independently of the practice itself, we shall fail. Property has been defined in terms of agreement, but the agreement (i.e. the *content* of the agreement) can itself only be defined by reference to the notion of property. The content and the object of the agreement are defined in terms of one another, and so we are going round in circles. There is no way to rationalise or justify the pattern of behaviour without circularity. This logical circle derives from the fact that the whole discourse, the whole language game of calling something 'property', is a self-referring practice. In virtue of it being a self-referring practice it is also a self-creating practice.

Let us test these ideas against our remaining example: the institution of marriage. Being married is a social status not a physical state. A close inspection of an individual will not reveal whether they are married, except indirectly and contingently (e.g. if they are wearing a wedding ring). Nor will such an inspection reveal what their being married consists in because that property isn't intrinsic to their person, or even to their mind. Following the previous analyses we may say being married consists in their being *called* 'married' by other people. This is an over-simplification, and some of the complexities will be introduced in the next chapter, but it is on the right lines. We can approach it in two steps. First, a couple are married if they are 'called' married by someone who has the authority to make them married by so calling them (e.g. a priest or official). The second step is to ask what this authority consists in? Clearly, 'authority' is a social status, so if our previous simplifications are allowed, we can say this too rests on consensus. Someone is correctly called an authority because everyone else calls them this. If no one accepted their authority they would possess no such

characteristic. And here too we can raise the question of what someone is being called, when they are called an 'authority'. The substance, as it were, of this description must come down to patterns of behaviour, but once again its meaning cannot be articulated without going round in circles. If their being an authority is constituted by being so called, the content and the object of this 'calling' are locked together by circular definitions. Which is to say, the entire process once again shows itself to be self-referring and self-creating.

A circle of talk about talk, where the reference of the talk is the practice of reference itself, creates an immediate puzzle about its origins. It is like a juggling act: we marvel at how it could be set in motion. Barnes called this the problem of 'priming' the system. It is a feature of models of this kind that a separate account needs to be given of its origins, because the origins of the system must be understood in terms lying outside the system itself. Consider money again. In order to solve the problem of priming, some economists, speculating about the origins of money, insist something can only become a form of currency if it was initially felt to be intrinsically valuable. That is, most people found themselves independently and instinctively drawn to it, so when a market came to exist the commodity was readily disposable. This is why gold and silver provided the basis for many systems of coinage. No monetary system, they say, could have originated by an 'agreement', or act of legislation, that some arbitrarily chosen substance (e.g. pieces of paper) was to count as money (cf. Menger 1892).

Providing that some external trigger or stimulus can be found to 'prime' the system, social institutions can be analysed as self-referring systems of talk and thought.[4] If this seems strange, two analogies may make it more acceptable by throwing into relief different aspects of these collective, self-referential practices. First, consider the utterance, 'I greet you'. To say this *is* to greet someone – just as saying 'I welcome you' is to welcome them, or 'I curse you' is to curse them. To use Wittgenstein's own example of a funeral oration: when we say 'we mourn our . . . ' this is 'an expression of mourning, not to tell anything to those who are present' (PI II: p.189). These cases have been called 'performative utterances' (Austin 1961). A performative utterance makes itself true by being uttered (cf. Bach 1975). This gives us a simple way to sum up the foregoing analysis of social institutions. We can treat them like giant performative utterances, produced by the social collective. A

second analogy may add substance, and some internal structure, to this overall picture. Recall the phenomenon of the self-fulfilling prophecy that has caused so much trouble to social scientists wanting to make predictions. The word gets round that party A is tipped to win the election, and because of a band-wagon effect some voters shift their allegiance from party B and vote for party A – thus helping to make the prediction come true. Or the in-group say the out-group are idle, so they refuse to employ them, thus giving rise to an enforced idleness that is read as proof of the original claim. The model of institutions as self-referring and self-creating tells us that institutions are themselves a species of self-fulfilling prophecy (Krishna 1971).

We now have a simple answer to our question: what is an institution? It is a collective pattern of self-referring activity. This, in essence, is the idea proposed independently by Anscombe and Barnes. I will adopt it in what follows.[5] The task now is to locate these self-referential and performative processes within rules and rule following. Their most obvious and important role is in connection with the normative aspect of rules. The 'right' continuation, say, of a number series (which also defines what is meant by the rule) is that continuation which is collectively called 'right'. This is not a matter of counting up votes, but refers to a stable pattern of interaction. I call this continuation right because others call it right, but I am correct in calling it right on this basis because their calling it right makes it right. The rightness to which I refer is constituted by their references to its rightness, and my reference, of course, contributes to the phenomenon to which all the other speakers in their turn refer. The self-referential model explains how the rule itself is part of the currency of interaction, and a medium of self-understanding. The acts of reference to the rule are occasioned by commenting on the performances of others, and of one's self. Thus the rule 'exists' in and through the practice of citing it and invoking it in the course of training, in the course of enjoining others to follow it, and in the course of telling them they have not followed it, or not followed it correctly. All of these things are said to others and to oneself, and are heard being said by others. In standard sociological parlance, the rule is an 'actor's category'. It is not just a spectator's description of a group's behaviour, or an idea utilised by an outside theorist wishing to summarise and predict their behaviour. It is used by the actors themselves in such a way that the phenomenon of following a rule is not distinct from the

descriptions given of it. Rather, it is a body of behaviour which includes the behaviour of ascribing or withholding that very description itself.[6]

Wittgenstein didn't use the label 'actor's category' but he did say:

> For if you give me a description of how people are trained in following a rule and how they react correctly to the training, you will yourself employ the expression of a rule in the description and will presuppose that I understand it (RFM VII: 26).

The self-referential model makes sense of this claim and explains our inability to give a description of rule following activity without using descriptions of the kind the actors would themselves use, i.e. actors' categories. The model can also cast light in this way on the more cryptic, but repeated claim: 'The "must" shows that he has gone in a circle' (RFM VI: 7 and 8). This means, in our terms, that invoking normative categories is itself an act of participation in a self-referential practice having no independent source of justification.[7] Any attempt to justify the 'must' will only lead us back to the practice itself. In the light of the self-referential model of institutions we can see why Wittgenstein said it made no sense to imagine a rule, such as following a sign-post, being followed just once in the history of mankind. Without a multitude of applications, by a multitude of different people, there would be no self-creating system of reference of the kind constituting an institution. The model also sheds light on the requirement that the instances of any putative rule-following practice must be outer, observable actions, not things enacted inside our heads. As Wittgenstein put it: 'And if it is now said: "Isn't it enough for there to be an imaginary application?" the answer is: No' (RFM VI: 32).

The answer is no, because if an institution is a self-referring practice, the object of the talk, namely that which provides the reality to which it refers, is one and the same with the acts of reference. There is simply nothing of the relevant nature there at all, unless we specify a set of outer acts of self-reference. Outer actions are necessary if we are to have interaction, and without interactions we have no institutions.

Though I have had to reach outside Wittgenstein's own writings to find a definition of 'social institution' we can see where the self-referential model makes contact with Wittgenstein's text. Here is another example. In the middle of a discussion of mathematical

certainty he suddenly invoked a comparison with one of the classic cases of a self-fulfilling prophecy – the bank that breaks because of a run on the bank, brought about simply because people believe it will break. Wittgenstein asked:

> What sort of certainty is it that is based on the fact that in general there *won't* actually be a run on the banks by all their customers; though they would break if it did happen? (RFM VII: 35)

Characteristically he didn't wait to answer his own question, but there can be no doubt as to the correct thing to say. The certainty in question is not just a statistical certainty that withdrawals will be offset by deposits because random fluctuations will balance out. The certainty, in so far as there is certainty, is a belief in the systematic character of the beliefs of others to the effect that the bank is sound. Each person who believes in the soundness of the bank is believing something about a reality constituted, at least in part, by the totality of beliefs, all of which are about the bank's soundness, and hence about the system of belief itself. The phenomenon called 'soundness' is self-referring and self-creating (or, at least, has a major component of this character). However 'fundamentally' sound we may take a bank to be, if it is widely believed not to be sound then it truly isn't sound. Soundness is to a great extent made up of beliefs about soundness. Wittgenstein therefore posed his question in terms that *invited* the answer provided by the self-reference model.

The self-referential, performative model of an institution also makes sense of the following claim: 'what the correct following of a rule consists in cannot be described *more closely* than by describing the *learning* of "proceeding according to the rule"' (RFM VII: 26).

Given a description of a material object or physical or biological process we can, in principle, study it more closely by examining it more minutely or experimenting on it in more detail. This is because it has a material existence independent of our current descriptions and the current state of belief about it. A 'social object', by contrast, is constituted by the descriptions actors and participants give it. It has no existence independent of their beliefs and utterances about it; hence it cannot be described 'more closely' by, as it were, getting behind these descriptions. Because they are self-referring there is nothing behind them, and as I have argued, this is precisely the case with the normativity of rules.

Reconstructing Wittgenstein's argument in these terms also sheds light on one of his most puzzling and challenging claims about mathematics and mathematical logic. In his *Remarks on the Foundations of Mathematics* he criticised the programme, associated with Russell and Frege, of turning ordinary, everyday informal mathematics into a formal, deductive system or 'calculus'. The aim of this programme was to make mathematical reasoning more rigorous and secure. Turning informal reasoning into formal reasoning was thought to reduce the chances of slipping, unawares, into a contradiction. The next step was to try to prove these formal deductive systems themselves did not contain implicit and unnoticed contradictions. The same worry then recurred about the logical or mathematical apparatus used to construct the proofs of logical consistency. Wittgenstein viewed the entire exercise with profound suspicion. He was not critical of its technicalities – that was just more mathematics – but of the motives driving it along and the interpretation put on its technical achievements. Mathematicians and logicians spoke as if, without proofs of consistency, they were sleep-walking along a road between abysses they might fall down at any moment. In reply to this Wittgenstein defined his position as follows: 'Can we be certain that there are not abysses now that we do not see? But suppose I were to say: The abysses in a calculus are not there if I don't see them!' (RFM III: 78).

Such a position was, and remains, highly controversial (cf. Chihara 1982). It offends against the intuition that mathematics accords with some genuine and objective, independent reality. If mathematical systems do correspond to such a reality, then we would indeed be in danger of our beliefs about it getting out of alignment with that reality. The 'abysses' would be real. But if mathematics has the ontological status of an institution, then it falls within the scope of 'linguistic idealism': it is a reality having no existence independent of our collective thoughts about it, and references to it. Any individual can fall into the abyss of non-alignment with others, but there is no question of alignment or non-alignment for the institution as such.

Wittgenstein developed a similar theme in a discussion of the multiplication of 12 and 12. What, he asked, if we always went wrong when we multiplied 12 x 12? Would such a supposition even make sense? For those who think of mathematical reality as an independent realm of truth it surely must make sense. So suppose we were always wrong:

> True, it is unintelligible how this can have happened, but it has happened. So everything worked out in this way is wrong! – But what does it matter? It does not matter at all! – And in that case there must be something wrong in our idea of the truth and falsity of arithmetical propositions (RFM I: 135).

The voice framing the reply – saying that it would not matter at all if we always went wrong – is Wittgenstein's own. He located the error in the idea of the truth and falsity of a mathematical proposition. The error – we can now see – is to think mathematical truth involves some form of correspondence between a mathematical statement and an independent mathematical reality. The correct analysis is one where it does not matter if everything is 'wrong', because then nothing would be wrong. This must be because, whatever it is that 'everything' comes to, is constitutive of the definition of truth. The usual idea of correspondence, the one Wittgenstein said was wrong, fails to take account of the self-fulfilling character of mathematics as an institutional practice. This is why he could say 'There *can* be no mistake of calculation in "12 x 12 = 144"' (RFM III: 73), just as there can be no mistake in saying the rules of chess are what they are collectively agreed to be.

An even simpler case is 2 x 2 = 4. We are tempted to say: but even if everybody thought that 2 x 2 = 5 it would still be four (PI II: p. 226). For Wittgenstein this was as fallacious as saying: people sometimes make mistakes in a game, so it is possible that in some games everybody makes nothing but false moves. It is an example of misunderstanding the logic of our expressions (PI: 345). A game in which everybody made such and such a move, where everyone regarded it as correct, would be a game in which the move really was correct. The game is constituted by the practice, and has no existence outside the practice. Games, for Wittgenstein, provide a simple, though supremely important, example of institutions. They provide his basic metaphor for language, so it is highly significant that they exemplify the self-reference model. What, then, would it be like if everyone believed 2 + 2 = 5? Wittgenstein's reply is interesting:

> Well, I could imagine, for instance, that people had a different calculus, or a technique which we should not call 'calculating'. But would it be *wrong*? (Is a coronation *wrong*? To beings different from ourselves it might look extremely odd.) (PII: p. 226–27).

Notice the comparison at the end with a coronation. A coronation is a collective act of deeming someone to be a monarch: it is a monarch-making episode. It is a creative performance. Someone becomes a monarch because the coronation makes them a monarch. What is it to *be* a monarch? Going back to our model, and imagining the simplest possible social characterisation of the role, we can say: to be a monarch is for everyone to acknowledge that person as a monarch. To be a monarch is, in our extended sense, to be called a monarch. So the implied answer to Wittgenstein's question, 'Is a coronation *wrong*?' is, of course, negative. It can't be wrong any more than it can be wrong to treat as money what everybody treats as money, or regard as property what everybody regards as property, or deem to be married those whom everybody deems as married. It can't be wrong because it is an institution, and institutions are self-referring, self-creating, and therefore self-validating.

Can an 'idealist' approach, such as this, really do justice to our central example, the rule for counting in twos? I have already shown how the self-referential model can easily cope with the normative aspect of this rule, but there are other problems to overcome. As we have just seen, if we adopt an idealist approach we are saying the reality in question has no existence outside human beliefs and practices. This clashes with our intuitions about the objectivity of the number sequence 2, 4, 6, 8. We might accept that a social status is created by acts of reference to it. We might even say this of a right, but can we swallow it for our number sequence? Any tendency to think of the rules of addition, or numbers, or sequences of numbers, as independent objects of discourse, works against the plausibility of the self-referential model.

The conviction that mathematical statements refer to an independent reality has, traditionally, expressed itself in two ways, giving rise to two different pictures of mathematical objectivity. One of these is usually called 'platonism' (or 'mathematical realism'). On this view we assume the existence of a special, abstract realm to house the objects referred to in mathematical propositions. The other picture, called 'empiricism', is adopted by those who want to identify the independent reality of mathematics with the ordinary, material world of everyday experience (Kitcher 1984). A platonist would say the sequence generated by the rule '+2' is an abstract but real object, or an abstract, but real, pattern of such objects, perceived by a special kind of intellectual intuition. The empiricist, by contrast, would say that such objects, or patterns of objects,

don't need any form of intuition, other than that provided by our sense-organs. For the empiricist, the sequence 2, 4, 6, 8, etc. can be exemplified by a suitably arranged row of, say, pebbles. These are independent and material enough to meet anyone's demands for realism. Platonists, such as Frege, viewed the empiricist's 'gingerbread or pebble arithmetic' with contempt, saying that it failed to do justice to the true objectivity of arithmetic (Frege 1884: VII; Bloor 1991: Chapter 4). But whatever the internal differences between these two approaches, we can see them as both providing a candidate reality at odds with an idealist account of the number sequence.

The same general sense of unease with the sociological analysis can be expressed in another way. Suppose we are following the rule '+2' by laying out sequences of pebbles. How do we find out if we are doing it correctly? Clearly, part of the answer is that we look carefully at the pebbles themselves. We attend to the precise characteristics inherent in their arrangement. That we look at the pebbles, rather than at other people, seems to tell against the idealist model. The pebbles constitute an independent object of just the kind missing in the case of, say, a right. Isn't it because we have access to some independent, non-social, reality that different rule followers can pursue the rule quite separately, and then come together and find themselves in agreement? The success of non-collusive co-ordination, as it can be called, implies access to an independent and shared reality by which to steer ourselves.

The important truth underlying these anti-idealist arguments is that mathematical concepts are modelled on empirical concepts, and mathematical operations are modelled on empirical operations. The prototype of addition is the physical bringing together of things. The idea that mathematicians or logicians can produce proofs of elementary additions, like 2 + 2 = 4, at some deeper and more basic level than the empirical, is an illusion. (They simply manipulate and count their symbols as the empiricist manipulates and counts the pebbles.) Empirical processes underlie the logical ones, not vice versa (Bloor 1994; Mackie 1966). Similarly, the prototype of subtraction and division is physical removal and partitioning, which is how we learned these processes at school. The manipulation of algebraic symbols is likewise grounded in the routines for ordering and sorting objects – and the story can be developed for more sophisticated concepts, with physical displacements and rotations underlying the concepts of vector, imaginary

number and group. Wittgenstein devoted much of the detailed discussion in his *Remarks on the Foundations of Mathematics* to exploring these themes, hence his claim that, 'Proof, one might say, must originally be a kind of experiment . . . ' (RFM III: 23). Adopting a variant of the empiricist theory of arithmetic removes much of the sting from the realist or anti-idealist argument. It explains how it is we can find the rule for '+2' exemplified in the non-social, material world, rather than solely in people's reactions to, or talk about, that world. It provides (along with the appeal to instinct) the resources for explaining non-collusive co-ordination. (For a general account of Wittgenstein's approach to mathematics see Bloor 1983: Chapter 5).

This still leaves a number of questions to be addressed. Why do some institutions (like 'rights') and some social statuses (like 'marriage') have no direct counterpart in the material world, whereas on the present account, arithmetical concepts do have such counterparts? Why can we lay out on the ground, for all to see, a 2, 4, 6, 8 sequence of pebbles, when we can't display any corresponding physical thing as the visible prototype of a right or the state of matrimony? A display of force accompanying the seizure of an object would not be the display of a right to it. Nor would a scene of domestic bliss be an empirical manifestation of the status of being married. If rights and marriage and number sequences are all examples of institutions, why the difference?

The problem can be pushed to a deeper level. The class of men is, to a first approximation, and leaving finitist complexities aside, a 'natural class'. It represents what we take to be a biological class to be found in nature. The class of married men, however, does not filter out any correspondingly natural sub-class, such as blue-eyed men, or fair-haired men. In general, social categories do not map onto our natural categories. They constitute arbitrary sub-classes of a natural kind, or a mixture of natural kinds.[8] How can this be reconciled with the idea that physical things are the prototypes of mathematical things, and empirical processes the models of logical ones? It still looks as if the attempt to accommodate the criticisms of the 'realists' is pulling us away from the sociological approach. The answer is that the physical objects which might exemplify the 2, 4, 6, 8 sequence do not, or need not, any more constitute a natural kind than do married men. The sequence can be created using pebbles, or humans, or gingerbreads, or any mixture. We must not confuse the object *qua* object, with the object *qua* member of the

sequence. That would be like confusing the coin *qua* coin, with the coin *qua* metal disc. In other words, rules may be structured around physical regularities, or embodied in them, and this is all the empirical model theory requires, but the two things are still qualitatively distinct. Similarly, rules may be structured around our instinctive tendencies to generate and discriminate regularities in our actions, but that does not make rules one and the same with these instincts.

Anyone developing an account of arithmetic as grounded in, and modelled on, experience, needs to attend carefully to these distinctions. This explains Wittgenstein's preoccupation with the difference between experiment and calculation. He called them the poles between which human activities move (RFM VII: 30) and tried to show how that polarity was constructed. He wanted to do justice to the realistic motive behind the identification of calculation and experiment but, at the same time as rejecting it, avoid the obscurantism of many of those who resist the assimilation (RFM III: 76). In outline his argument was this: calculation isn't experiment, for the simple and decisive reason that we don't treat it as experiment (RFM I: 161–165, III: 67–77). Propositions that could be treated as empirical propositions are, under some circumstances, turned from facts into standards (RFM I: 165). They are, he said metaphorically, deposited in the archives of language (RFM III: 29).

> It is as if we had hardened the empirical proposition into a rule. And now we have, not an hypothesis that gets tested by experience, but a paradigm with which experience is compared and judged. And so a new kind of judgement (RFM VI: 22).

Wittgenstein's preoccupation with the 'institution of measuring' (RFM III: 36) arises out of this concern. We measure the length of a plank of wood by seeing how many times another piece of wood – a yardstick – can be placed end-to-end along its side. This is a physical process involving the juxtaposition of physical objects. The aim is to see what happens, but it isn't an experiment. Similarly, the yardstick is a physical thing with physical properties, but it also has properties not locatable in its physical substance, but only locatable around it, in the way it is treated. It has the property of being accorded a certain use, and a certain status by a group of people. It is a standard length because it is used as a standard length by them. If they call it a standard, it is a standard. Mathematical propositions, Wittgenstein implied, are parts of an institution in the same way a standard measure of length is part of an institution (RFM

III: 36). The qualitative difference between mathematical and empirical propositions, the 'new kind of judgement' of which he spoke, derives from the fact that 'mathematics as such is always measure, not thing measured' (RFM III: 75). Platonists are sensitive to this difference, hence their refusal to follow the empiricists and identify mathematical reality with empirical reality. They rightly see the number one is not to be identified with 'a thing' (in the ordinary sense of 'thing') and that a mathematical rule is not a law of nature of a highly general kind. Unfortunately, in their response, they over-react and lose grip on reality. Wittgenstein saw this mistake for what it was: a reification, or subliming, of social processes (RFM V: 16; cf. PI: 38, 89, 94).

The task Wittgenstein set himself, of saying what is right and what is wrong with the assimilation of calculation to experiment, of doing justice to the 'realism' of empiricism without embracing its errors, was a difficult one. He characterised the difficulty with a striking metaphor, picking up the theme already touched on above in the image of the coronation. He spoke of our feeling that 'mathematics stands on a pedestal' (RFM VII: 6). A rule, he said, is something we detach from experience, so that it 'stands as it were alone in its glory', though it is the facts of daily experience that give it its importance (RFM VII: 3).

> What I have to do is something like describing the office of a king; – in doing which I must never fall into the error of explaining the kingly dignity by the king's usefulness, but I must leave neither his usefulness nor his dignity out of account (RFM VII: 3).

From the standpoint of the self-referential model of social institutions, we can recognise the significance of this figure of speech. It sums up the interplay of the social and non-social we find in arithmetical phenomena such as our rule-governed number sequence.

I have now provided a simple model of a social institution. It goes beyond what Wittgenstein had to say, but it makes contact with his discussion at a number of significant points. The main feature of the model is its self-referential character. These collective processes of self-reference leave their mark at the individual, psychological level. In the next chapter I shall take the discussion back to the structure and content of the states of mind of the rule follower.

4

CONSCIENTIOUSNESS

A distinction is often drawn between genuinely *following* a rule and merely behaving in ways which outwardly *conform* to it. In the first case the outer behaviour derives from the rule; in the second it derives from some other source but happens, by coincidence, to fit the specifications. This is an interesting distinction and, surely, a correct one. We need to mark this difference in our practical affairs. I want to look at the ways this distinction bears on the idea of rules as institutions, and how it connects with the development of Wittgenstein's sociological approach outlined in the previous chapter.

The distinction between following and conforming is most often and most conspicuously employed in connection with moral rules or principles. Adherents to the Kantian tradition in ethics have long insisted that a truly moral act must derive from what they call a 'reverence' for the moral law (Paton 1946: Chapter V). If we help those in need because of a gush of sympathy, rather than because of a moral imperative to do so, then (they say) our act is not a truly moral one. This applies however beneficial the action might be in its consequences. Critics who see this as cold and forbidding miss the point. Feelings are by their very nature variable and at the mercy of the preoccupations of a busy life. Those to whom we have a genuine obligation should not have to wait on our inclination or convenience. The Kantians are right: duty is independent of subjective feelings and circumstances. A moral act – that is, one deriving from a moral rule – must involve a conscientious concern with the pure morality of the matter.

As well as this heavy-duty formulation of the distinction, there is an equally effective light-weight version. Long before rigorists talked of the categorical imperative, irreverent street urchins would

have been making a version of the same point. 'Get out!', orders the voice of authority; 'I was going anyway', comes the reply. This is not only impertinent, it is also philosophically astute. If you are going anyway (i.e. going of your own volition and in pursuit of your own purposes), then you are not going because of the order. An order to do X is only obeyed if X is done, and done only because of the order and with the intention of obeying it. So the urchin prudently does what is required, but does not admit to doing it because it is required, thereby artfully and deliberately accomplishing the moral shortfall against which the Kantians warned us. The distinction between following and conforming does not, therefore, only exist in the books of the philosophers, it is part of the currency of everyday interaction. It is clearly applicable, not only to orders, but to all cases of rule following, from number sequences to the Ten Commandments. (As Wittgenstein said, following a rule is analogous to obeying an order (PI: 206)). A rule is followed if, and only if, the actors bring about the conformity of their behaviour with the rule by intending to follow it. They must, as it were, have the rule before their mind as their guide and goal. We may call this the conscientiousness condition.

Wittgenstein's acceptance of the conscientiousness condition is revealed in a number of ways. In a passage I will return to shortly, he said that making a move in chess doesn't just consist in moving a chess piece in certain ways on the board (PI: 33). That would be mere conformity. Nor, he said, does obeying an order simply mean responding to what is said in the way the speaker wants. The exercise is only one of giving and obeying orders if the effect comes about through the process of obeying, not through some other process. Thus:

> When I say that the orders 'Bring me sugar' and 'Bring me milk' make sense, but not the combination 'Milk me sugar', that does not mean that the utterance of this combination of words has no effect. And if its effect is that the other person stares at me and gapes, I don't on that account call it the order to stare and gape, even if that was precisely the effect that I wanted to produce (PI: 498).

We don't here have the distinction of following and conforming, or the conscientiousness condition, explicitly drawn, but we are given observations which illustrate and, so to speak, etch out the distinction.

Perhaps Wittgenstein didn't offer an explicit formulation of the conscientiousness condition because he wasn't convinced of the accuracy of any simple and general statement of the requirement. Groups differ in the extent to which they concern themselves with the inner state of the agent rather than the outer form of their actions, and circumstances affect cases. If all our rule-bound doings were rigidly subject to a conscientiousness condition we could not 'accidentally' break a rule, and yet unintentional violations are an accepted category. Rules governing various forms of 'pollution' (e.g. in the rituals of food preparation) focus on the outer behaviour, not the knowledge or intentions of the agent. Again, not knowing what he was doing didn't help Oedipus. (Gilbert Murray brings this case under the rubric of pollution in the preface to his translation – Murray 1911: vi). There are systematic cultural variables to be explored. In his *A Short History of Ethics*, MacIntyre (1967) has traced the historical shift from concrete, behavioural, role-specific concepts of virtue to more abstract, personal, and subjective ones. The conscientiousness condition, therefore, is not without its complexities. In what follows I shall put such problems aside and concentrate on the logical skeleton of certain simplified cases.

Even in its simplest form, however, the distinction between following and conforming has some puzzling consequences. It also poses a problem for Wittgenstein's theory. Let us say that to obey an order you must (at least) *think* you are obeying it. Obedience without the requisite accompaniment of thought would be mere conformity. Similarly, to make a move in a game of chess the players must (at least) *think* they are making a move, thus ruling out the non-player who accidentally nudges a chess piece without noticing it. Someone who goes through the moves of a game without realising what they are about, or without having the requisite understanding of its status or point, isn't really playing the game at all. They are doing something else, and what they are doing depends in part on what they think they are doing. By the same token, a genuine rule-follower must at least *think* he is following the rule. The general formula for this class of actions, i.e. those requiring the conscientiousness condition, is: X-ing implies *thinking* you are X-ing.

Anscombe (1969) has pointed out the remarkable fact that, in the class of cases identified by this formula, it is impossible to explain what it is to engage in the activity of X-ing. This impossibility derives from the built-in circularity infecting any attempt at

explanation. Using her own example, assume getting married involves thinking you are getting married. (If you think its just a rehearsal that invalidates the proceedings.) To explain getting married involves mentioning thinking you are getting married, but explaining thinking you are getting married involves mentioning getting married. This takes us round in a circle. Clearly, not much by way of 'thinking' is required to go through with a marriage, but that doesn't invalidate Anscombe's point. However little is deemed necessary, the circle will still be generated. It doesn't depend on making unrealistic demands about, say, the legal knowledge of the participants. Stating the argument generally, we can say: to explain what it is to X requires we mention thinking you are X-ing, because this is one of the essential ingredients of X-ing. We must then explain what it is to *think* you are X-ing. Explaining this thought means explaining its content, and that can only be done by mentioning X-ing – which is what we set out to explain in the first place. The attempt at explanation gets us nowhere. This argument clearly applies to rule following. To follow a rule involves thinking you are following a rule, so to explain rule following we must mention thinking you are following a rule. To explain 'thinking you are following a rule' we must explain what the thought is about, which takes us back to rule following, which was our starting point.

Our seemingly evident and simple conscientiousness condition, with its Kantian pedigree and its street credibility, seems to be in trouble. Perhaps what seems evident is really incoherent, and there is something wrong with the conscientiousness condition. Working in the opposite direction, however, we can argue that we do have a coherent conception of following a rule; so there must be ways of circumventing the circularity problem without giving up the conscientiousness condition. Here is a conjecture about how we accomplish a practical solution to the problem. We acquire the ideas of 'following a rule' and 'thinking we are following a rule' at the same time as a connected pair, with all their inter-connections given along with them. We acquire a conceptual scheme with the circle built into it, so we do not need to break into the circle to explain it. 'Light' as Wittgenstein charmingly expressed it, 'dawns gradually over the whole' (OC: 141). There is nothing unusual in this. All it means is we acquire these concepts as we acquire the concepts of nut and bolt, master and slave, male and female, and husband and wife. A nut is something that threads onto a bolt; and a bolt is something nuts thread onto, and so on for the other corre-

sponding and inter-defined pairs. We learn the pairs by being introduced to activities where both are on display in relation to one another, and both are spoken of in the course of the activity.

In the case of rule following we learn to pair 'following' and 'thinking you are following' but *not* by logical definition or verbal explanation. Perhaps such 'definitions' can be given after a fashion, but they are no use here (cf. RFM VI: 18). We learn by being *socialised* into the practice called 'following a rule'. We are shown how to go on, we are started off and stopped, guided, given examples, and encouraged by gestures and words. We are taught the appropriate verbal responses to thread into the activity. If, in the course of this socialisation, our teachers tried to articulate and justify the relevant ideas in a step-by-step fashion, they would find they went round in circles. At the basic level, however, teaching does not proceed by explanation, and hence escapes the sequential and linear requirements of definition. It is concrete and holistic, exploiting a form of trust that does not require each step to be justified in turn, and where the learner is prepared to understand the earlier in terms of the later.

The circularity problem, then, is circumvented in practice by the process of socialisation into an existing set of rule-following practices. Suppose we are learning a game, and this learning is at such an early stage of individual development that it cannot rely on a background of accumulated experiences of other games. We do not and cannot have the concept of a 'move' in the game defined for us, but we can watch and try to take part and respond to sanctions. In this way most people are able to become competent performers in what is, for them, a first game. Acquiring a modicum of competence will obviously involve our hearing and seeing and handling and attending to many things. It will involve our mobilising our memory and organising our actions in order to sustain the participation. We will, in short, naturally come to 'think' we are playing the game in the course of playing it, and we will learn to call the requisite mental orientation 'thinking' or 'intending' such and such. This story presupposes an established practice. To be socialised is to be socialised into an existing custom or institution. How do such processes get established? How can we *create* an activity whose component thoughts embrace and presuppose the activity itself without raising the spectre of circularity again? The problem of creating a conscientious practice is one we have met before. It was introduced in the last chapter under the rubric of 'the problem of

priming'. The discussion of socialisation, especially when offered as a solution to the problem of circularity, clearly presupposes a solution to the problem of how a practice can ever get started. The problem of specifying the requirement of conscientiousness, in a manner which makes it amenable to a solution, is just a version of the more general problem of priming. This problematic feature of rules derives from their sociological nature. Whenever we have institutions of the kind described by the self-reference model, we will have knowledgeable and conscientious agents whose mental states are inexplicable except by reference to the practices in which they participate – and whose practices are inexplicable except by reference to their mental states.

Wittgenstein's treatment of these themes is suggestive, but disconnected and in need of filling out. His idea was that the requisite inner states of the rule follower gain their identity, and the content imputed to them, through the circumstances or surroundings in which they operate. In a passage from which a brief quotation has already been taken he said:

> a move in chess doesn't consist simply in moving a piece in such-and-such a way on the board – nor yet in one's thoughts and feelings as one makes the move: but in the circumstances that we call 'playing a game of chess', 'solving a chess problem', and so on (PI: 33).

A 'move', therefore, is not simply (a) the outer behaviour, nor simply (b) thoughts and feelings. What about the combination of the two, the outer move combined with the relevant thoughts? If these thoughts were to the effect that the outer performance was indeed a move in chess, wouldn't that suffice? According to our simplified schema that X-ing implies thinking you are X-ing, the answer is 'yes'. Nevertheless, Wittgenstein insisted, the secret of what makes the action into a 'move' lies in 'circumstances'. This is what connects the inner and outer parts together, because the content of the inner thoughts itself depends on these circumstances. Hence he said: 'If the technique of the game of chess did not exist, I could not intend to play a game of chess' (PI: 337). So the inner element of X-ing – namely, thinking you are X-ing – could not itself exist without the technique known as X-ing. Why is this? Wittgenstein's answer again invoked the surrounding circumstances. An intention, he said, is 'embedded in its situation, in human customs and institutions' (PI: 337). In making this claim he was

deliberately negating the idea that thoughts and beliefs and intentions are fully definable in a purely internal, mental or personal way, i.e. in a way abstracted from background and social context (cf. PI: 205). At the beginning of the *Investigations* he emphasised the contextual principle by an analogy drawn from engineering:

> 'I set the brake up by connecting up rod and lever' – Yes, given the whole of the rest of the mechanism. Only in conjunction with that is it a brake-lever, and separated from its support it is not even a lever; it may be anything or nothing (PI: 6).

If we have thoughts that accompany and connect to our actions, then they are what they are, and have the meaning they do, in virtue of the wider context of action.

Wittgenstein took the argument one stage further. In a difficult paragraph he linked two themes: first, the role of the context in giving meaning to our mental states, and second, the performative and self-referencing process by which the context is itself made up. The case under discussion was that of an order: someone is told to leave the room. We have already seen that to obey the order the leaving must be done because of the order, and with the purpose of obeying it. What manner of link is that? Are we necessarily aware if it comes into play? This has always troubled persons of a conscientious bent. How do they know they are really being moral? Might they think they were acting out of respect for morality, when really it was from unworthy motives? This threatens to play havoc with the conscientiousness condition and the distinction between following and conforming. Wittgenstein touched on this question with his example of the order to leave the room. He considered the two assertions: 'I am leaving the room because you tell me to' and 'I am leaving the room, but not because you tell me to'. He then asked:

> Does this proposition *describe* a connexion between my action and his order; or does it make the connexion?
>
> Can one ask: 'How do you know that you do it because of this, or not because of this?' And is the answer perhaps: 'I feel it' (PI: 487)?

A similar question is broached later at PI: 682: 'is he observing that the connexion existed, or is he making it by means of these words?' Here the problem was not mixed motives, but selective attention and discriminating the reference of words. Someone says, 'it will stop soon'. Did he mean the pain he was feeling, or the sound of the

piano from the next room? If we ask him and he tells us, is he reporting a bond that already exists, or is he creating that bond? Wittgenstein replied: 'Can't I say *both*?' This is a less than satisfactory reply, and needed to be explained at greater length. Wittgenstein's point was that certain features of the situation, 'didn't simply make their appearance with the words' (PI: 684). These features will be the various tendencies and dispositions and feelings of the speaker that underlie the utterance. These are not created by the words, and are in some sense there to be 'observed' or reported on by the speaker. On the other hand, the words, 'I meant the piano, not the pain', do add something new to the situation. Whatever their causes they make a link in the public realm others can orient to, notice, refer to, repeat, cite in justification or criticism at a later date, use as evidence about the speaker's mind or character or sensitivities, and so on. We might say it represents a contribution to a developing and continuing conversation. Any particular contribution of this kind might be trivial, but in the light of our self-referential model we can see Wittgenstein was putting his finger on processes that can be highly consequential – as his own question about the soundness of the bank reminds us. In their immediate context, however, Wittgenstein raised these issues in order to shed light on the identity of certain mental states: was the state of mind directed at the order? Was it directed at the sound, or the pain? He was drawing our attention to the contextual and performative contributions to that identity. The position to which Wittgenstein was tending could be summed up as follows: the classification of mental states is like the classification of social objects or statuses. Or, even more simply: mental states *are* social states.

These conclusions are derived from hints rather than developed themes in Wittgenstein's writing. Nevertheless, the hints are intriguing, and fit in well with the self-referring model of institutions described in the last chapter. Usually, however, Wittgenstein didn't touch on the creation of the social context of our utterances, but took it as a stable, taken-for-granted background, available to do the job of contextualisation. When he said the game of chess confers identity on our thoughts about chess and the intention to play chess, he did not discuss how these conventions are constituted or created. Rather, we were given an account telling us that 'thinking about playing chess' (i.e. the conscientiousness condition) simply amounts to our being a normal and competent person, who has been exposed to chess playing practices, and has acquired the

capacity to participate. The outer and the inner aspects of rule following are created and connected together in the course of socialisation (PI: 197, cf. Anscombe 1969).

There is a second problem arising out of conscientiousness. If the first was that it seemed to make rule following inexplicable, the second is that it stands in apparent contradiction to one of Wittgenstein's most significant conclusions. Recall his saying that we cannot always follow rules by interpreting the rule to ourselves; consequently at some level I must 'obey the rule *blindly*' (PI: 219). Now the problem is this: if following a rule involves thinking you are following a rule, then it seems that you cannot do this 'blindly'. Conversely, if you are following a rule blindly, this would seem to preclude 'thinking'. Either Wittgenstein was wrong in his analysis of the role of interpretation, or the distinction between following and conforming is called into question. The two requirements can only be reconciled if we can make sense of 'blind conscientiousness', which sounds wrong.

Both the problem and its resolution can be better understood if we take the discussion back to the moral case. We surely do not want, or need, to preclude the possibility that a truly moral response can be automatic. Indeed, it is something like an ideal that we be so thoroughly moral that it comes naturally to respond in accordance with the requirements of morality. It is also part of common-sense psychology to see that there are dangers here: the person who no longer has to struggle with temptation may lapse into complacency. Nevertheless, we should not confuse what has become 'blind', in the sense of habitual, with what is 'blind' in the sense of being entirely devoid of thought or awareness. We can take it that it was the former to which Wittgenstein referred when he said we follow a rule blindly. And it is clearly the latter that would prevent our differentiating 'following', as a conscientious process, from mere conformity. Wittgenstein drew the relevant distinction as follows: 'One follows the rule *mechanically*. Hence one compares it with a mechanism. . . . "Mechanical" – that means: without thinking. But *entirely* without thinking? Without *reflecting*' (RFM VII: 60).

'Blind' or 'mechanical' rule following, then, is not entirely 'without thinking'. It is automatic in the sense of needing no *reflection*. When we say following a rule involves thinking you are following it, this needs to be interpreted minimally. All the 'thinking' required is the routine awareness of the average, competent member

of a society who has been socialised into its customs and institutions. Wittgenstein sometimes expressed this by saying what was in question was a matter of 'attitude'. 'Our children are not only given practice in calculation but are also trained to adopt a particular attitude towards a mistake in calculating' (RFM VII: 61).

In a footnoted variant of this passage the phrase 'mistake in calculating' is replaced by 'departure from the norm' (cf. Z: 299). What is the attitude in question? Elsewhere Wittgenstein explained it in what might be called phenomenological terms. Someone who calculates in 'my language-game' does not feel it is a peculiarity of *their* nature that they get the outcome they do. They are under the impression of following a thread that is already there. They act in a matter of course way and only know one explanation for their action, namely 'how the thread goes'.

> He does just let himself go on when he follows the rule or the examples; however, he does not regard what he does as a peculiarity of *his* course; he says, not: 'so *that's* how I went', but: 'so *that's* how it goes' (RFM VII: 4).

Here, then, is Wittgenstein's reconciliation of the blind character of rule following with the conscientiousness condition. We can be 'blindly conscientious'. We follow some rules automatically, but do so within a social framework to which we are known to be responsive, and within which we operate according to acceptable standards of competence and awareness. In this way we can be said to 'think' or 'know' that we are following a rule, even though we are responding 'blindly'. The conscientiousness condition is not, ultimately, at odds with what Wittgenstein wants to say about interpretation coming to an end.

There is one final topic concerning conscientiousness which needs to be examined. The conscientiousness requirement says: X-ing implies thinking you are X-ing. Thinking you are following a rule is a necessary condition for following a rule. What about the claim that it is a sufficient condition? Are there circumstances where thinking you are X-ing implies you *are* X-ing? In particular, are there any cases of rule following where thinking you are obeying the rule means you *are* obeying it? Wittgenstein said no, but we should reflect on the following historical example where this very principle received practical expression in the conduct of life. There are people for whom the voice of conscience is everything. They sometimes declare themselves guided by an inner light, where that inner light is

infallible. It is infallible because it comes from God. God reveals Himself directly to each person, because each person participates in the divine nature. Notions of part and whole and divisibility do not apply to God, so we cannot say one part of Him is in the soul of each creature; we must think of His presence in each case as absolute and complete. Though one may hesitate to express it like this, the individual believer listening to the voice of conscience becomes, or *is*, God. Whatever moves the conscience of the believer is infallibly right, and its dictates stand infinitely higher than the demands of one's fellows, or of transient and superficial things like the law of the land. If conscience were to bid us kill, then killing would be right. To think it is right is for it to be right.

Strange though it may sound, this position has had many adherents. In a simplified form it describes the theological and ideological doctrines of the radical Protestant sects of seventeenth-century Europe. In England these were mainly recruited from the betrayed remnants of Cromwell's army and the 'masterless men' who roamed the country looking for work and means of survival after the disruption of the civil wars. Their antinomian ideology of subjectivism and radical individualism was condemned by their contemporaries under the name of 'enthusiasm'. These were the people who, in a famous phrase of the time, threatened to 'turn the world upside down'. They refused to pay taxes, interrupted sermons, and demanded the right to preach as the spirit moved them. For them, no church institutions or clerical hierarchy should interpose itself between the worshipper and God. This relationship was to be direct, simple and unmediated. They claimed back the common land for the common people, and declared their right to a share in the fruits of the earth God had created in common for everyone. They flouted convention and sexual morality, and disregarded the property rights of the great landowners. They felt moved by the divine force within them, and sometimes made claims to be Christ. They were called, and called themselves, by such names as the Levellers, the Diggers, the Ranters and the Quakers. The individuals involved ranged from the impressive Gerard Winstanley, the leader of the Diggers or True Levellers, to the bizarre and unfortunate James Nayler who, in 1656, entered Bristol on a donkey, proclaiming his divinity. These men and women were widely perceived as a threat to social order, denounced by respectable persons, refuted by respectable intellectuals, and hunted by the authorities who killed or imprisoned and tortured them (Hill 1975).

One of the most famous philosophical polemics against the sectaries was John Locke's 'On Enthusiasm', added to the 1699 edition of his *An Essay Concerning Human Understanding*. Locke saw these people as a menace, and identified the root of their doctrinal errors in their failure to distinguish being right from believing they were right. Since they assumed the voice speaking to them in their minds was the voice of God, they felt they could do no wrong as long as they followed their supposed revelation. This reasoning, said Locke, led them in a circle: 'It is a revelation, because they firmly believe it; and they firmly believe it, because it is a revelation' (Locke *Essay* 1699: IV.19.10). We can only get out of this circle if we insist that our convictions be subject to an external standard – which Locke called 'reason'. Thus:

> Every conceit that thoroughly warms our fancies must pass for an inspiration, if there be nothing but the strength of our persuasions, whereby to judge of our persuasions. If reason must not examine their truth by something extrinsical to the persuasions themselves, inspirations and delusions, truth and falsehood, will have the same measure, and will not be possible to be distinguished.
>
> (Locke *Essay* 1699: IV.19.14)

This old dispute over the role of private judgement in theology and ethics (which was at the same time a full-blooded political conflict) makes contact with Wittgenstein's text at a number of points and in rather intriguing ways. I am thinking of those passages in which he contrasted the following of rules with responding to an inspiration, and also his discussion of how a rule might be thought to 'intimate' to us how we are to proceed by means of an 'inner voice' or 'a voice within me' (PI: 222–237). If rule following were really like this – that is, a case of harkening, or some kind of receptivity – then we could not require someone else to follow the rule in the same way as we do (PI: 232). In short, such a theory cannot be right, because it makes nonsense of normativity: it simply sets us off along the road that leads to the conclusion Locke identified and denounced. Wittgenstein's famous polemic against a 'private language' – that is, a language whose words refer to immediate private sensations, a language of pure subjectivity informed by purely private, inner rules – can be seen as just another tract in the anti-subjectivist, anti-individualist tradition of Locke's attack on 'enthusiasm' (cf. PI:

243–270). The resonance with Locke's passage is clear when Wittgenstein wrote:

> And hence also 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it (PI: 202).

For Wittgenstein you cannot create a rule just by thinking you have done so. To justify behaviour by reference to a rule existing simply in the imagination is not enough. Subjective justification was no more acceptable to him than it was to Locke and his contemporaries denouncing the Levellers and Ranters and Quakers: 'justification consists in appealing to something independent' (PI: 265). To believe in a rule is one thing; for there really to be a rule is another. What is it for there to *be* a rule? Wittgenstein's answer was clear. The authentic, independent or 'extrinsical', existence of a rule – independent or extrinsic, that is, to any individual consciousness – consists in its being a social institution.[1]

How cogent are the considerations advanced by Locke against the enthusiast, and by Wittgenstein against the would-be private linguist? The answer is that neither of their arguments can claim absolute cogency or a decisive victory. Their claims will seem evidently right to some, and just as evidently wrong to others and there is nothing to compel either side to change. Take Locke's version of the argument first. He confronted the individualist with a sharply drawn picture of his conduct, and he has spoken in a way that would set many heads nodding in agreement, but he has done nothing to compel the enthusiast to abandon his claims. The enthusiast just knows in his heart, and that conviction is invincible. He believes – and no mockery about his merely believing that he believes will force him to budge (cf. PI: 260). Such force as Locke's argument possessed did not derive from its logical properties, but from its sociological bearings. Locke was trading on the alleged consequences of enthusiasm, namely its threat to the established social order. He was saying, in effect: this will lead to anarchy. To those, like Locke, who opposed it, this anarchy spelled the end of all social order; to those who embraced it, it represented an alternative order.

The same has to be said about Wittgenstein's attack on the idea of private rule following and a subjective language of inner experience. Wittgenstein could no more refute the private linguist than

Locke could refute the enthusiast. All he could do was to point to the consequences of holding this view and seek to encourage us in certain habits of description and classification, e.g. in the way we use the word 'language'. His impotence in this regard may be illustrated by one of his own analogies. In the course of his argument against private, inner languages he asked : why can't my right hand give my left hand a gift of money? Plausibly, he took it for granted that it can't because, 'the further practical consequences would not be those of a gift' (PI: 268). The entire procedure would be an idle ceremony, like giving a private definition of a word in the recesses of the mind by directing one's attention at a private sensation and inwardly undertaking to use a sign to designate it (cf. PI: 262, 268). But what if someone were unmoved by these considerations, and declared that it seemed wholly reasonable to them to bring the transaction between the right and left hand under the rubric of gift giving? To the argument about further practical consequences they could reply: true, this act of giving doesn't have all the usual consequences, but it is still a case of giving. (Perhaps it has other, inner consequences of comparable importance). Nothing Wittgenstein could say could budge such a person, nor on Wittgenstein's own deepest principles should it. All he could do is to point to certain consequences and deplore them.[2]

Any attack on individualism must try to convey a sense of the social changes and losses involved in adopting an individualist policy. (This is why the argument will inevitably have a variable persuasive power.) Both Locke and Wittgenstein were reminding us of what we would lack if we did not live within a social group. The real significance of Locke's philosophical argument was political, that of Wittgenstein's is methodological. It directs us to treat certain cases of what we call 'rule following' as paradigmatic, and others as peripheral and degenerate. Call the performance involving the right hand and the left hand 'gift giving' if you must, but don't think it will illuminate the social institution of the gift or the social significance of the activity. To say this is as good as saying: don't think it will illuminate gift giving. And what holds for gifts holds for rules. Such reminders are necessary because individualists are prone to forget, or fail to see such things. Thus a recent critic of the collectivist approach to rules correctly characterised its appeal to consensus as a form of honouring certain responses and dishonouring others. 'The community . . . has a practice of dignifying its members as saying things correctly or incorrectly, and in the light of

this practice it says that all its members do mean the same.' Unfortunately, the critic had prefaced this by saying: 'We do not know what a community would be lacking if its members failed to see each other in this way' (Blackburn 1984: 293). Collectivists, he went on, 'gave no account of what a community would be missing if it failed to see itself in this light' (ibid.: 299). This charge represents an extraordinary failure of imagination. The answer is that a community which did not see itself in this light would be missing, not merely one of its institutions, but the basis of all its institutions and the preconditions of social organisation.[3]

I have now completed my exposition and development of Wittgenstein's theory of rule following, and it is time to connect it to the existing literature. I shall use this as a testing ground for my interpretation of Wittgenstein's text and the theories I have extracted from it and added to it. This exercise will not take us away from the phenomenon of rule following itself, but provides a forum in which the subject can be examined from a variety of different angles.

5

RULE SCEPTICISM

As an example of rule following consider the task of adding together two numbers. If we know how to add 7 and 5 to make 12 we can say that, thus far, we know how to follow 'the rules of addition'. To follow the rules we must know the words 'add' and 'plus' mean a specific mathematical operation or function, and that if we encounter the signs 7 + 5, the '+' also means add or plus. We intend these words and signs to mean add or plus. Let us home-in on this process of meaning addition. What is it to mean a certain thing by a word we use, or by a thought we have? How do we *mean* things? It sounds like something we do, and presumably do with great frequency, so what is it that we do when we do it?

In his *Wittgenstein on Rules and Private Language* (1982), Saul Kripke characterised Wittgenstein's argument as a sceptical challenge directed at the claim that we can perform the feat of meaning addition by 'plus', or indeed that we can perform the feat of meaning anything by our words. He said we can understand Wittgenstein as posing the question: what fact is it about yourself that constitutes your meaning addition by 'plus'? We might, on occasion, say it is a fact about Jones that he means such and such by a word he uses. In which case, what is this fact? We ought to be able to specify it. Wittgenstein's argument (as Kripke presents it) is designed to show that, try as we may, we cannot produce any such fact. No fact of any kind corresponds to talk about people meaning things. But if no fact corresponds to it, wouldn't that imply our talk about meaning things by our words was itself meaningless? The 'no fact' claim seems to threaten the notion of meaning in its entirety, which is surely an unacceptable result. Something must therefore be wrong with the terms of the discussion and indeed, Kripke goes on, there is something wrong: it rests on a false supposition. The chal-

lenge, and the failed attempt to respond to it, both presuppose our talk about meaning this or that by our words has to be of a fact-stating kind. The entire exchange has gone ahead on the assumption that the business of our 'meaning talk' is to state some fact or other. The challenge 'show me what fact it is about yourself that constitutes your meaning addition' presupposes there *is* some fact, and there *ought* to be such a fact. Kripke says that Wittgenstein held this presupposition to be in error. It may be true that we can't meet the sceptic's challenge to produce the relevant fact, and the sceptic may be right to doubt there is such a fact, but the failure may be of no great moment. Our discourse about meaning may be of an inherently non fact-stating kind. We can therefore grant the sceptic his case, but argue that no disconcerting consequences follow from the admission.

What is the character of our talk about meaning something by our words or thoughts if it doesn't state facts? Wittgenstein's answer to this, says Kripke, is to be understood against the background of his overall approach to language. On this approach, to say a proposition is true does not mean it 'corresponds' to an independent reality. Instead, calling a proposition true means it meets some generally understood condition of acceptability, e.g. that this verbal formulation is accepted as proper in these circumstances, that it would generally be deemed justified, that this is the conventional thing to say, or the traditional or convenient or useful thing. These provide what can be called the 'assertability conditions' for the sentence. Instead of thinking of our sentences, on whatever topic, as representing an independent reality, we should think of them instrumentally and pragmatically. We must see our talk of 'Jones means plus' in this light. Our failure to find the fact corresponding to this imputation will then not disturb us. The conditions under which it is proper to say Jones means addition by 'plus' will be, for example, those that justify our predicting he would answer '12' if asked to compute 7 plus 5. Those conditions are that Jones has undergone training in addition, and has shown himself to be competent. To impute a specific meaning to his words under these conditions is part of slotting him into our plans and interactions. Such imputations implicate and are implicated by the trust, co-operation and predictability which play such an important role in our lives. The imputation doesn't mean, directly, that he will slot in or can be trusted. It signalises this conditionally and indirectly, because it relates to just one strand in a complicated pattern of

interaction. Nevertheless, it gives us a right to trust Jones, and a right to reproach him and accuse him of error or incompetence if our expectations are thwarted, and interactions with him founder. In other words, the imputation tells us what he ought to do rather than precisely what he will do.[1]

This, in outline, is the argument Kripke attributes to Wittgenstein. We have a 'sceptical challenge' to the factual character of meaning, that is, a challenge to produce the fact corresponding to our talk about meaning. We are then given a 'sceptical solution' to the challenge, i.e. one based on the acknowledgement that no such fact can be located. Accepting the sceptical solution amounts to side-stepping this difficulty by insisting that the real significance of our meaning talk can be explained by reference to the role of imputations of meaning in our daily lives. These imputations mediate our interactions with one another. The substance of these interactions provide the conditions under which it is deemed appropriate to talk about 'meanings', i.e. they provide the assertion conditions for such talk. Clearly the assertion condition approach is, in a sense, a sociological account of meaning. It is one of the available ways of construing the role of meaning talk in the organisation of our collective affairs. There is, therefore, a similarity between Kripke's reading and the one advanced in the previous chapters, but the two accounts are not the same. The full extent of the divergence will become clear as we probe more deeply into Kripke's version of Wittgenstein's argument.

Suppose a sceptic decides to cause trouble by asking us to add, say, 68 and 57. We reply 125, but the sceptic makes the astonishing claim that our answer should be 5. He says in the past, when we used the word 'add', we really meant, not addition, but another mathematical function he calls 'quaddition'. Quaddition is the same as addition up to a point, but then diverges from it. Quaddition is defined as follows:

$$a \oplus b = a + b \qquad \text{if } a, b, < 57$$
$$= 5 \qquad \text{otherwise}$$

where $a \oplus b$ is read as 'a quadd b'. To understand this strange argument we must realise that in the past I can only have done a finite number of addition sums. There must be some so large that I have never encountered them. To keep things simple, and the numbers small, suppose 68 + 57 is one of these. The supposition we must

make is that the sceptic has cunningly asked us to perform a novel sum. In answering his question we are about to do a sum we have never done before. We follow normal routines and give our answer confidently, only to be told this is not the answer we ought to have given. The sceptic insists that to be consistent with our past intentions we should have said 5. Naturally we will want to reply that we know what we mean and what we have meant in the past, and it never so much as occurred to us to mean quadd. This is the point at which the sceptic politely asks to be shown the fact that constitutes our meaning add. He is not, he assures us, trying to convince us we have a bad memory. He will allow us to assume we have a perfect memory for what we meant, but he wants to know what fact it is about ourselves we are remembering with such clarity.

Kripke surveys for us some of the more obvious, candidate facts that might be offered in answer to the sceptic. They are the following. First, we have the fact of our past practice. Surely all the addition sums we have done in the past attest to our meaning add? Unfortunately they don't, says the sceptic, because our past practice could just as well be identified as quadding, because of the overlap of the two functions. If our past practice is, as the sceptic says, seen as quadding, then of course we should now be saying '5' not '125'. So the fact of our past practice doesn't differentiate our having added from our having quadded and cannot therefore furnish the fact constituting our meaning plus by 'add'.

Second, suppose we cite not facts about our outer behaviour, but inner facts about our feelings, e.g. the feeling or qualitative experience of adding. When we responded to 68 + 57 with '125', presumably it felt like the old, familiar experience of adding. Won't this show that we were meaning to add? Clearly not, because feelings are really neither here nor there. Even if it were true that in the past adding always had a characteristic feel to it, this would at most accompany and not constitute the meaning. Feelings cannot provide the facts we are looking for.

Third, and more plausibly, what about our dispositions? At school our training will have given us certain habits, response tendencies and dispositions to act in certain ways. Surely, despite its supposed novelty, these dispositions would be triggered when we were asked to give the sum of 68 and 57? Tempting though it is, the sceptic has two good reasons to reject this suggestion. To begin with, the addition function is infinite, while our dispositions, like our life, are finite. Nobody could literally have a disposition to add *any* two

numbers, or to do *all* addition sums, and yet that is, surely, the scope of the word 'plus'. This gap between our finite dispositions and the infinite character of the addition function is enough to preclude their identification.[2] Even if we were able to convince ourselves, and the sceptic, that our dispositions can be viewed as potentially infinite, a further and even deeper reason can be advanced against the disposition theory. We often have dispositions to make mistakes when we add, e.g. habitually forgetting to carry a number. We treat such dispositions as things which lead us to deviate from what we mean to do. They are the cause of our making mistakes, thus showing that what we mean can't be identified with what we are disposed to do. If whatever we were disposed to do were read back into our intentions, then we would lose our grip on the notion of making a mistake, i.e. on normativity. In general, as Kripke says, to identify a disposition is to say what we are inclined to do, not what we should do. So dispositions can't furnish the fact of meaning that would enable us to give a straight answer to the sceptic.

Fourth, suppose when we gave the answer to the sum 68 plus 57 we were following a set of self-instructions, and these were explicitly instructions on how to add. Wouldn't *that* prove we were adding? Suppose we talked ourselves through the following procedure: to add x and y, take a large heap of pebbles and count out a sub-heap of x pebbles; then count out from the remains of the large heap a separate sub-heap of y pebbles; then bring the two sub-heaps together and count them – that will give the sum of x and y. If *that* is what we were doing with 68 and 57, wouldn't it prove we were adding and not quadding? Again, the sceptic has a cogent reply. He began by raising a problem about the meaning of 'addition', so can't he now raise exactly the same kind of problem about all the words in the set of instructions that were meant to tell us how to add? Trying to cure a non-standard reading of addition, by giving a rule for adding, only invites the sceptic to repeat his challenge on the new level, that is, by turning his scepticism on the rule for following the rule. For instance, how do we know that when we instructed ourselves to 'count' we had not in the past, all along, meant 'quount'? Quounting, the sceptic will explain, is just like counting – until you get to 57.

By this stage there seems to be only one move open to us with which to reply to the sceptic. We have no option but to be dogmatic and insist that meaning is a special kind of fact in its own right. It is unanalysable and irreducible to other terms. It can't be identified

with any *other* fact about us, because it is unique. The sceptic will rightly say this is just an evasion, leaving the nature of meaning as mysterious as it was at the outset. We can now see the disturbing force of the sceptical challenge. Unless we are prepared to be dogmatic and treat meaning as irreducible, we have so far failed to produce the fact that constitutes our meaning addition – or our meaning anything at all. The prospect looks bleak – unless we reject the fact stating character of our talk about meaning and our imputations of meaning. According to Kripke, this is exactly what Wittgenstein did.

Having now described Kripke's argument (i.e. Wittgenstein's argument as Kripke presents it) I want to make some criticisms of it. I take myself to be criticising the exposition rather than the underlying argument in Wittgenstein's text, though at the end of the chapter I shall add one small qualification to that. Let us, then, look at the steps of the sceptical argument again. What precisely has the sceptical challenge established? Has it really established that no fact at all can constitute meaning? Or has it established that no individualistic fact can play this role, i.e. no fact about an individual's behaviour, or inner feelings, mental states, dispositions or verbal self-instructions? Unfortunately Kripke himself poses the challenge in an equivocal way, sometimes presenting it as the challenge to produce *any* fact, sometimes challenging us to produce a fact about what is in the mind or brain of the individual. The equivocation is easily illustrated. Starting with the stronger claim, on p.13 of his book, the sceptical position is said to be that 'there can . . . be no fact about which function I meant'. This is repeated on p. 39, where the sceptic 'claims that an omniscient being, with access to *all* available facts, still would not find any fact that differentiates between the plus and the quus hypotheses'. Similarly, a few pages later, Kripke refers to the sceptic's 'denial that there is any fact of the matter as to what I meant' (k: 41). Again, 'There are . . . no truth conditions or facts in virtue of which it can be the case that he accords with his past intentions or not' (k: 89). These quotations all point to the sweeping 'no fact thesis', but other formulations introduce significant qualifications. For example, on p.11 we are told that to meet the sceptical challenge head on, we 'must give an account of what fact it is (about my mental state) that constitutes my meaning plus'. Sometimes the sceptic is presented as saying, 'nothing in my mental history' establishes whether I mean plus or quus (k: 21). Notice the change: the claim here is not that no fact about me

determines what I mean, but that no fact about my internal, mental state will do this job. And perhaps most explicitly of all, we are told that what is really denied is: 'that the notion of a person following a given rule is to be analyzed simply in terms of facts about the rule follower and the rule follower alone, without reference to his membership in a wider community' (K: 109).

It is clear from these illustrations – and others could be provided – that Kripke's text oscillates between presenting the sceptical challenge as (1) a challenge to the factual character of meaning, and (2) a challenge to the individualistic character of meaning. Which is it? The answer is obvious from an inspection of the steps of the sceptical argument itself as they were set out above. The candidates presented for consideration as the fact of meaning, and which are successfully rebutted and shown to be inadequate, are all *individualistic* facts. All along, the real challenge is to individualistic accounts of meaning. The sceptic has challenged the individualist to produce an individualistic fact – a fact, say, about our minds; and the sceptic has won, because the challenge has not been met in a plausible or adequate way. It is only in so far as we share the individualist's prejudices that we will jump from the result that there are no individualistic facts of meaning to the conclusion that there are no facts of meaning at all. Unfortunately, this is just the step taken in the rest of the argument as Kripke presents it. The conclusion is drawn that the sceptic has established there are no facts of meaning, and this is why it is necessary to launch a rescue operation by claiming that discourse about meaning is not of a fact-stating kind. This is why assertion conditions, rather than correspondence conditions, are brought in to complete the story and to provide a candidate solution to the problem posed by the sceptic. But what about the possibility of producing a sociological fact, i.e. a fact about society and the agent's place in it, and relation to it? By taking the drastic step of leaving fact-stating discourse behind entirely, and constructing an account based on assertion conditions, the full potential for a straight sociological answer is never properly explored. He confronts the possibility of such an answer but quickly dismisses it as, 'a social, or community-wide, version of the disposition theory' (K: 111), that is, as an enlarged version of an approach already shown to have failed. I shall return to this point later, but to prepare the ground for a more sympathetic response to the straight, sociological answer to the sceptic consider the following analogy.

Suppose we say Jones owns so many acres of land. What sort of

claim is this? Clearly it is about Jones. We take ourselves to be asserting a fact about him – namely, the fact of his ownership. We now encounter Kripke's sceptic, flushed with his victory in the field of mathematics. He thinks legal facts will be easy meat after arithmetic. He therefore challenges us to produce the fact of the matter constituting Jones' ownership. Obviously, if we inspect Jones' person, or his mind or his behaviour in isolation, we shall fail to meet the challenge. That is to say, if we look for an individualistic fact we will not find the fact of ownership; but then, few of us would expect to find it in those places. We should look elsewhere, to the contracts Jones had signed, and to the deeds in his name. We need to look around Jones, not at him, and locate the relevant legal institutions that define his status as a property owner.

To drive the point home, let us replay the moves of the sceptical argument. First, would Jones' past practice establish or constitute his ownership? Would the fact that he has ploughed his acres these last forty years clinch the matter? Clearly not, because he might have been exploiting the land without a right to do so. Second, do his feelings yield the relevant facts? For instance, he may have feelings of familiarity and rootedness in the land. Obviously, feelings are irrelevant, because he might have such feelings about land that is not his, and fail to have them for land that is, without this in any way impinging on his legal standing as an owner. Third, do his dispositions give him any title? The answer is no, and would still be no even if there were dispositions characteristic of ownership, e.g. the disposition to put a fence round the land and to chase off intruders. What are we to do? Should we postulate a unique, non-natural, irreducible and indefinable fact about Jones that constitutes his state of being-an-owner? Clearly, this is intellectually disreputable and unscientific. Are we to draw the general conclusion that ownership of land is not a fact at all? If this means anything it means there is no such thing as owning land and, taken at face value, this is glaringly false.

The correct response is to see that the sceptical argument, as it has been rehearsed here, only calls into question the individualistic character of ownership, not its factual character as such. Ownership can be true of Jones, even if it is not true because of Jones, that is, Jones alone. What kind of truth is a truth about ownership? Clearly it is a truth within the framework of the institution of property. It is a matter of fact, but the fact in question is a social fact. 'Ownership' can be attributed to an individual in virtue

of their place within a social system in which people organise their affairs around the concept of property. This will be a system in which people can speak of the 'rights' of ownership. If the sceptic were to press the question, 'what do the words in this practice refer to?', the answer, ultimately, would have to be: to other instances of the same referring practice. Because we are dealing with a social institution, it has the form of a self-referring system. It is a collective performative which creates such things as property rights by virtue of referring to them. At its simplest, something is correctly called 'Jones' land' because everybody, or nearly everybody, is disposed to call it 'Jones' land', and shares an idea of what it is for something to 'belong' to somebody.

The example of 'Jones owns such and such number of acres' makes it clear why we might be tempted to postulate a *mysterious* fact to answer the sceptic. At least we might be tempted if we are individualists. Convinced that we *must* be able to find an individualistic fact, but possessing a dim awareness of the social reality that in truth constitutes the relevant fact, we misconstrue the social fact, treating it as if it were an individualistic fact. We compress the complex surroundings into a characteristic of Jones himself, and then react by finding it a very strange characteristic indeed. Hence we get the confused idea of an irreducible fact with peculiar mental or spiritual qualities. The error of seeing the social through the distorting lens of individualism explains the desperate character of some of the moves we might be tempted to make to defend our mistaken starting point.

An institutional analysis of the nature of property is not meant as a substitute for what Jones could or should say as he manoeuvres within the legal structure. Like Wittgenstein's own talk of language games it is best thought of, in the first instance, as formulated from a standpoint outside the language game, rather than being a move within it. The institutional theory tells us, in general terms, about the factual causes and preconditions for talk within the game possessing the meaning it does. It addresses the reality behind it. We must not conflate (a) truth conditions as they are formulated within the language game with (b) the facts of the matter as they pertain to the existence and character of the game itself. This could be expressed by saying that we must not confuse games and meta-games, though Wittgenstein himself always rejected this imagery of 'meta' levels (see Shanker 1988). There is good reason to share Wittgenstein's reservations here. The metaphors of 'game' and

'meta-game', or of the 'inside' and 'outside' of a language game, need to be used with caution. For good sociological reasons they can only be provisional and relative distinctions. The boundaries of language games, like the boundaries of concepts themselves, are social accomplishments. They too have the status of conventions or institutions. There is nothing, in principle, stopping Jones (or somebody he is chasing off his land) making the effort to bring in higher order, philosophical or sociological theories into their discourse about first order legal issues. Usually this will be futile, just as it is futile to play the game of Kripke's sceptic outside the philosophy seminar. This, however, is a contingent matter about how groups conduct their affairs. We cannot assume *a priori* that reflections originating 'outside' a language game will either leave it untouched, or will necessarily tend to undermine it. It can go either way depending on how the players of the games choose to deploy these resources. Abstruse doctrines can deliberately be used, positively or negatively, as vehicles for the expression of social interests.

How can we put this comparison to use and move from, say, 'Jones owns five acres' to 'Jones means addition'? We must see the state of meaning something as like the state of owning something. It is an institutional fact about an individual – that is, a statement about their membership of, or role in, an institution and a set of ongoing practices. That step has already been taken, and it was the purpose of Chapter 3 to explain it. The conclusion there was that our saying 'He is following a rule' has to be analysed in terms of the actor's participation in the activity that is called 'following a rule'. That participation comes from (a) being trained up in the relevant practice, (b) interacting with other participants, and (c) sharing their vocabulary and referential practices. Similarly with 'Jones means addition'. The (finite) content of the state of meaning, or the act of meaning, derives from the actor's real or perceived contact with the institution of adding. The institution itself only exists in virtue of the whole nexus of similar actions and references and behaviours by the other participants.

How does this institutional account differ from the assertion condition approach used by Kripke in his reconstruction of Wittgenstein's argument? The main difference is that the institutional account is based on the process of self-reference, and this finds no acknowledgement in the assertion condition analysis. There is no doubt that a model of language in terms of assertion conditions can capture many features of how we talk, and resonates

with much Wittgenstein has said. His later work was informed by a strong tendency not to take propositions at their face value. It is full of locutions like 'here we are inclined to say', 'this is what we say', 'we are inclined to use this metaphor', and so on.[3] At one point he even characterises his general position in terms of the insight that we should avoid confusing 'is' and 'is called' (RFM I: 127). Nevertheless, presenting us with a choice between correspondence and assertion conditions is not satisfactory. The dichotomy is false because it overlooks a third possibility. Between reference to an independent reality, and having no reference at all, we have self-reference, i.e. reference to a reality, but a reality which is dependent on the very acts of reference that are directed at it. This situation is realised when a group develops a self-referential practice, i.e. one whose terms refer to the practice itself. It was argued in Chapter 3 that this is the essence of a social institution. The assertion condition approach may be criticised on the grounds that it only captures a superficial aspect of this process, and neglects its central features.

In one of the most difficult and compressed parts of his exposition Kripke offers, on Wittgenstein's behalf, the following argument which seems to tell against the institutional approach and in favour of the assertion condition approach. If the appeal to institutions constitutes a straight answer to the sceptic, as it seems to, then it will surely have to provide truth conditions for our talk about meaning. Such theories, says Kripke, are open to many of the sceptical objections which hold against individualistic disposition theories.

In response it must be accepted that the institutional approach can, indeed, be considered as *a form* of community-wide dispositional theory. There is no reason to evade this fact. (Isn't the assertion condition approach also predicated on community-wide dispositions?) An institution can be looked upon as the collective product of the interactions between the dispositions of many individuals. A community-wide disposition theory, fashioned along the lines of the self-referring model, does not however fall into the traps the sceptic set for the individual disposition theory. The central difficulty for the attempt to equate meaning with an individual disposition is its failure to account for normativity. To be disposed is one thing; to be disposed rightly or wrongly is another. That was the sceptic's main point. By contrast an institution, which arises from the interaction of individual dispositions, does provide a normative basis for the actions of the individuals who are within it.

The important theme here is that of interaction. It is interaction, not mere aggregation, that makes a group a group. I shall later provide an example of the failure to appreciate this fact.

What about the sceptic's other observation, about finitude? An individual disposition to add is finite, while the addition function is infinite. Surely, a collection of individuals can still only muster a finite capacity however well-organised they are? This is true but not decisive. Recall the discussion of infinity and the finitist account of how operations with that concept can be understood. It is possible to have finite practices which sustain concepts with 'infinite' content. But surely, if the collectivist can appeal to finitism here, then so can the individualist. Unfortunately for the individualist this is not so, and the reason goes back to the basic issue of normativity. The individual's finite dispositions cannot sustain any conceptual practices, because concepts involve normativity, and this is precisely the commodity individuals cannot generate for themselves. This is why advocates of the individual version of the disposition theory cannot help themselves to the same replies as the collectivists.

If we cite a sociological fact in reply to the sceptic, can't this fact itself be rendered ambiguous between add and quadd, just as individual performances, or feelings, or dispositions, or self-instructions were rendered ambiguous? The difficulty here is more apparent than real. Consider a group of people, just like ourselves, who engage in adding. Just like us, they have the institution of addition. This amounts to saying that the things they do, and the practice they sustain, they call 'addition'. They train their children, as we train ours, and they also teach them that the practice they are being trained in is addition. Let us continue with Kripke's useful fiction and suppose that so far none of their addition sums has exceeded, let us say, 125. Now: is this *really* addition? What about the ambiguity in their practice between adding and quadding? The answer is: what ambiguity? There is no ambiguity. Within the practice so described there is no double meaning attached to the word 'add'. There is no surmise and no doubt (OC: 523–524). There is just one thing that they rightly do, namely give *this* answer to *this* question, and *that* answer to *that* question: '*This* is how one calculates. Calculation is *this*. What we learn at school, for example. Forget this transcendent certainty, which is connected with your concept of spirit' (OC: 47).

Ambiguity is ambiguity of sense, and sense is what is created by and in the practice. The 'sense' of a word can itself be called an

institution because it is created by the users of that word. It only exists in and through that collective use – which is why the individualist cannot make a comparable reply. There is no *closer* description of the sense of the word 'add' than belongs to the step-by-step practice of its users.

Meaning finitists say meaning is a moment by moment creation. A new situation may always arise in which routines are disrupted or change direction. This means: those engaged in the routine will define it as having been disrupted or changed. Confronted by such an awareness of change, concept users have a choice. They can either say they have changed the meaning of a term, or that, all along, its meaning has been stable but some aspects of past usage have been mistaken. From the standpoint of meaning finitism these conclusions are not discoveries but decisions. Such decisions can be made retrospective as we interpret and re-interpret the past. This is all part of the moment by moment creation of meaning. Now imagine someone in the community we are considering, constructing the quaddition function. They thereby create the possibility of the community declaring its own past practices ambiguous. This now becomes an option, and the community is, in principle, free to adopt it as part of their self-understanding. They are also free to reject it. They can say that never having thought of quaddition when they previously engaged in what they called 'addition', then there is no question of their having practised quaddition. As proof it wasn't quaddition they were doing, all that is needed is a collective endorsement of the answer 125, not 5, to the novel, ground-breaking addition sum of 68 + 57. If they deem 125 to be consistent with their past meaning, then it is consistent – because meaning is an institution, and an institution is what people make it. It truly is what they say it is, because it is constituted by what they say.

To those who do not accept meaning finitism this will look as if the members of our community are closing their eyes to a real possibility. The Wittgensteinian answer is simple: they are shut (PI II: p. 224). We must remember that for Wittgenstein the abysses in a logical system we do not see are simply not there at all: they do not exist. To shut our eyes in this sense, and under these circumstances, is not to ignore something really there, it is to remove it. This is an acceptable position precisely because the abysses in question are not abysses of rock, but of meaning. Meaning is something we create, not of course carelessly or frivolously, but collectively and pragmatically.

Although a properly formulated sociological approach prevents the sceptic from simply repeating the same argument to the same effect at the collective level, there is nevertheless some point to the reiterated sceptical challenge. It reminds us that the social collective is, in some respects, in an analogous position to the isolated individual, though not, of course, to the individual as a member of society. When building up the picture of meaning finitism I began with instincts. I said at this stage of the story, before social interaction had been introduced, we had no resources for explaining the normativity of rules and meanings. In the same way, the sociological theory clearly has no resources for explaining how a social collectivity, taken as a whole, can be right or wrong. The only possibility would be to put it in the context of other such collectivities, thus making a higher order collectivity; but then the same problem recurs at the next level. Ultimately there is no standard outside the social collectivity according to which it can be judged. Evaluative notions only make sense for individuals within a society. But if, according to meaning finitism, the group makes up the meanings of its concepts as it goes along, why can't the individual? The correct reply is that with an appropriately adjusted and normatively impoverished sense of 'concept' perhaps they can, but they cannot thereby reproduce the circumstances that apply to actions taken within a social context. The isolated individual and the isolated group are in an analogous position, in that neither of them can follow a rule, if 'rule' is to mean more than 'regularity'. Neither can operate according to a rule in the paradigmatic sense, because neither entity can bind itself by what it says or does. From a naturalistic perspective we must conclude that the imposition of a binding necessity is a within-group phenomenon – that is to say: a matter of social institutions, conventions and customs.

How do members of a linguistic community come to accept that they agree with one another, as they must if they are going to coordinate and concert their actions to maintain a consensus? How do they even arrive at the conviction that a candidate answer given to the sceptic was 125 rather than 5? The sceptical argument, it has been said, will eat into the sociological solution at this level too (cf. Hoffman 1985). This problem is really a version of Wittgenstein's own idea that anything follows from anything on some interpretation. The correct response must follow his own solution that our language games are built on a level of response which does not have the form of an interpretation. We must say the agreement involved

here is not mediated by meanings, but must be automatic. Odd though it may sound, it will be agreement without any propositional content, that is, agreement in actions (RFM VI: 39, cf. PI: 241).[4]

I have shown that Kripke's formulation contains an ambiguity, sometimes being presented as an argument against the factual character of meaning, sometimes against an individualist picture of meaning. It is only fair to point out that hints, though no more than hints, of the same ambiguity can also be detected in Wittgenstein himself. One side of the ambiguity occurs in connection with his discussion of double negation. He said it was possible to use the negation sign according to two different conventions. The one usually adopted by logicians says 'not not P' is equivalent to 'P'. Another convention might be that repetition did not cancel, but strengthened, the negation, making it more emphatic. Now, asked Wittgenstein, 'when I uttered the double negation, what constituted my meaning it as a strengthened negative and not as an affirmative' (PI: 557)?

He replied, in terms Kripke captured so well, 'There is no answer running: "It consisted in the fact that . . . " ' (PI: 557). This appears to be a statement of the unqualified 'no fact thesis'. No fact about me constitutes my meaning the sign in one of these ways rather than the other. Now look at Wittgenstein's discussion, in another part of the *Investigations,* of quite a different theme, namely the proposition, 'At that word we both thought of him'. (This is something we might say when reporting a conversation in which some turn of phrase had brought someone to our minds.) The question is: what went on when we both thought of him? Wittgenstein made a striking claim about such an event, though he couched it in metaphorical terms. He said, 'If God had looked into our minds he would not have been able to see there whom we were speaking of' (PI II: p. 217).

Presumably, God knows everything there is to know, so He certainly knows of whom we were thinking when we had the experience described above. Wittgenstein's point was not that He didn't know whom we were talking about, but that He could not know this by looking into our minds (and nor, of course, could He know by looking at our dispositions, for the good reasons set out by Kripke's sceptic). This is not the 'no fact' thesis, but the 'no individualistic fact' thesis. So what is Wittgenstein's position: that there are no facts of meaning, or no individualistic facts? The weight of evidence is clear. Wittgenstein's real target was individualism, not the factual

thesis as such. After all, his famous slogan that the meaning of a word is its use in the language (PI: 43) is telling us the kind of fact that constitutes meaning. The aim of his argument throughout the *Investigations* was to get us to shift our attention from mental processes, and individual psychological events, to patterns of interaction within social collectivities, i.e. to 'language games', conventions, customs, and institutions.

6

THE ANALOGY WITH VON MISES

Kripke concludes his account of the sceptical solution with a cautiously worded footnote, saying:

> it has struck me that there is perhaps a certain analogy between Wittgenstein's private language argument and Ludwig von Mises's celebrated argument concerning economic calculation under socialism (K: 112, fn 88).

Von Mises (1935, 1963) argued that under socialism economic calculation would be impossible. Rational choice requires an economic agent to compare prices, and prices are set by other people. So who sets the prices for the Central Planning Board of a socialist commonwealth? Ultimately, no one. Kripke glosses the non-market decisions of the central planners as a case where 'whatever seems to be right will be right', so we cannot talk about right. That is, the price is whatever they say it is, so we cannot talk about prices and, without prices, there can be no calculation. Kripke says it is his impression that, though von Mises' argument points to real problems for planners, it is 'now almost universally rejected as a theoretical proposition'.[1] He therefore wonders, 'whether the fact bodes at all ill' for Wittgenstein's account of rule following (K: 112).

Kripke is right to see an analogy here, but the comparison is with quite general features of economic systems, and is largely independent of von Mises' attack on socialism. I shall pursue the hint given by Kripke in order to lend support to my own 'straight' sociological answer to the sceptic. First, however, a certain oddity needs addressing. Doesn't the proposed analogy get things the wrong way round? Von Mises is attacking central administrative control, so in some sense of the word, he is an 'individualist'. Should not individualists feel sympathy towards private rules and private languages, as

a celebration of the autonomy of the individual mind? Surely, von Mises' position ought to be analogous to that of supporters of private languages and private rules, not analogous to Wittgenstein's anti-individualist attack on them.

The oddity is removed when we realise von Mises is not an individualist at all in certain important respects. He endows economic agents with subjective preferences, so he is an individualist to that extent. He argues by building up a picture of the interactions of such individuals. So he is an individualist in that respect too. But, he insists, to calculate rationally, agents need more than their subjective valuations. They need something objective to grasp, like prices, and these are the irreducible upshot of numerous bargains struck between individuals as they interact to form a market. That is where the individualism ends. Thus: 'Prices are . . . social phenomena as they are brought about by the interplay of the valuations of all individuals participating in the operation of the market' (HA: 331). Price is not, like preference, an individualistic fact, but a collective fact derived from a specific form of collective organisation.

Wittgenstein often used buying and selling as examples in his writing (e.g. PI: 1, 120, 142; RFM I: 143–150) but the deeper analogy between von Mises and Wittgenstein is that, for both thinkers, the sources of objective constraint in our thinking come from interaction with other people. Furthermore, the concepts which embody and define these constraints (e.g. 'price' or 'rule' or 'meaning') are themselves the medium of the interaction. Money and market prices have the character of institutions. They are created through activity informed by these very concepts themselves. Notice the self-referential character of von Mises' claim that: 'each separate calculation . . . depends exclusively on the fact that it is precisely in market dealings that market prices to be taken as the basis of calculation are formed . . . ' (1935: 111).

The calculation, he is telling us, depends on and refers to, a reality made up of just such calculations. This is not to deny that the calculations may refer to commodities of an entirely material nature, but it reminds us that a material object only becomes a commodity in the context of a market.

The felt objectivity and external, constraining power of a market price is not an absolute thing. Whilst it is genuinely external to us as individuals, it is only external in proportion to the negligible contribution that our own activity makes to it. Thus:

> Each individual, in buying or not buying, and in selling or not selling, contributes his share to the formation of the market prices. But the larger the market is, the smaller is the weight of each individual's contribution. Thus the structure of market prices appears to the individual as a datum to which he must adjust his conduct (HA: 331).

Von Mises' worry about centralisation was its dependence on an agency for whom prices were not external realities. Such an agency could only decide, not calculate. The underlying point, however, is not exclusive to command economies but applies equally to market processes. For a small shareholder on the stock market the price of a share at any given time can be treated as a simple, external fact. An agency with a giant holding, however, will find the price is influenced by its own decisions. If it decides to sell, that brings the price down, and vice versa. For our purposes it does not matter if von Mises' opposition to central planning hits its target or not. His argument is simply one way amongst others of emphasising the significance of self-referring, collective processes. (The same point could be made by using Keynes' discussion of speculation from the chapter on long-term expectation in his *General Theory*, see Keynes 1973: Chapter 12.)

The analogy touches on the basic tenets of finitism. An economic calculation is based upon present prices but what, asks von Mises, is a 'present price'? Prices are things people have actually paid – in the past. These historical transactions never come to an end in a 'final' price. Nor do they more or less accurately reflect a 'real price'. Talk of 'final prices', says von Mises, is strictly a mistake: it is 'merely a mental tool' to help us, as economic actors, deploy these historical precedents in forecasting (and hence constituting) future prices (HA: 332). Past prices do not determine future prices. Only the anticipations of future prices entertained by buyers and sellers can do that.

This is a 'finitist' theory of prices. Just as past transactions do not create a stable and efficacious thing called 'the present price', so past usage doesn't create a determinate thing called 'the present meaning' of a word, sign or belief. Just as prices, considered as an abstraction from past episodes, have no causal powers, so likewise meanings are causally inert. The past enters into the story, in both cases, only as a precedent. The notion of the 'real meaning' of a concept or a sign deserves the same scorn as economists reserve for

the outdated and unscientific notion of the 'real' or 'just' price of a commodity. The only real price is the price paid in the course of real transactions as they proceed *von Fall zu Fall*. There is no standard outside these transactions. Nor could a price be created by a single transaction. Not just once in the history of mankind could someone sell a commodity. Similarly the only basis for identifying meaning is by reference to the actual use of the sign in the course of a sequence of collective applications. What that meaning will come to be, in the future, is not determined by past applications. These yield at most a set of resources and contingencies and precedents.

Kripke's sceptical challenge is parallel to von Mises' sceptical challenge directed at the believers in the just price. Von Mises' analysis amounts to the challenge: 'point to the fact that constitutes that price'. All we can find, on or about the person of the individual entrepreneur, will be a certain history of past transactions, a certain disposition to buy or sell, and a certain set of valuations or subjective preferences. Each of these candidates can be ruled out in turn as the fact that constitutes the price. If we then claim to find in the mind of the entrepreneur a knowledge or a belief regarding a price, we are permitting ourselves to bring into the account the surrounding system of market relations, i.e. the very thing we were seeking to illuminate. The entrepreneurs' belief about past, present and future prices has no other identifiable reference than the system of market relations which are themselves mediated by this kind of belief. The only 'fact', then, is the fact of the system itself – the market taken as a whole. Concepts like 'price' are used to isolate and capture an aspect of this system, to identify a moment within this whole. As von Mises put it:

> The market process is coherent and indivisible. It is an indissoluble intertwinement of actions and reactions, of moves and countermoves. But the insufficiency of our mental abilities enjoins upon us the necessity of dividing it into parts and analysing each of these parts separately. In resorting to such artificial cleavages we must never forget that the seemingly autonomous existence of these parts is an imaginary makeshift of our minds. They are only parts, that is, they cannot even be thought of as existing outside the structure of which they are parts (HA: 333).

The parts of the social structure called a 'market' have no existence outside the whole. This declaration has a surprisingly holistic ring

for a self-professed individualist, but the force of the vision cannot be denied. What von Mises, in his own way and for his own ideological purposes, did for the special case of the market system, Wittgenstein did for our conceptual system generally. We must not reify meaning but keep it tied down to the context of use and interaction, then its finitist and social character will be evident.

The analogy between Wittgenstein's and von Mises' argument is clearly real. When investigated it reinforces the claim that the correct formulation of the sceptical solution needs an apparatus, not of assertion conditions, but of self-referring and performative processes. Furthermore, if we can talk about 'price' as a fact of economic life, and 'ownership' as a fact of legal life, then by the same principles we may take 'meaning' to be a fact of our collective life. In his final footnote on von Mises, Kripke unwittingly pointed to the 'straight', sociological solution to the sceptical challenge, a solution of the very kind he officially opposed.

7

INDIVIDUALISM

Colin McGinn's book *Wittgenstein on Meaning* (1984) is meant as an answer to Kripke's sceptical challenge, and a criticism of the sociological or collectivist position Kripke attributes to Wittgenstein. According to McGinn, Wittgenstein did not develop a sceptical argument about rules, and he did not endorse a sociological account of rule following, either in the form attributed to him by Kripke or in any variant of it. On the contrary, says McGinn, Wittgenstein was an individualist, and offered an individualist account of rule following. Although on occasion identifying shortcomings in Wittgenstein's treatment of these themes, McGinn endorses what he sees as Wittgenstein's individualist approach:

> My own view is that rule-following may be conceived, as I think it is pre-theoretically, in entirely individualistic terms. I should emphasise that I do not believe that Wittgenstein himself would have dissented from this conclusion: insofar as he has a view on the individual/social opposition, he is an individualist . . . (M: 200).

The general character of McGinn's discussion is well illustrated by one of the 'straight' replies to the sceptical challenge he canvasses. This reply is called the irreducibility thesis. The challenge, remember, is to say what fact it is about us that constitutes our meaning what we do, e.g. meaning add by 'plus'. A 'straight' reply offers some candidate fact. The basis of the irreducibility thesis is that the sceptic is playing a trick on us: distracting us from remembering that meaning something is a fact just as it stands (M: 151). McGinn makes much of the allegation that Kripke has smuggled an assumption into the discussion by tacitly ruling out this reply. Hasn't the sceptic got us to believe that the fact of meaning must be

identifiable as some other sort of fact, i.e. a fact that isn't at first sight a fact of meaning, but to which (allegedly) meaning can be reduced? Thus, says McGinn:

> The kind of reply that is being implicitly judged illicit is one that simply uses semantic concepts, as follows: what it consists in to mean/refer to addition by '+' is for the speaker to *mean/refer* to addition by '+' – *this* is the sort of 'fact' that meaning consists in (M: 151).

The aim here is clearly to push the burden of proof back onto the sceptic, and to demand that he say why we should not take 'semantic discourse' as fact-stating in its own right. But is it really Kripke's sceptic who, as McGinn puts it, 'needs to defend an undefended and undisclosed premise' (M: 151–152)? And if it is the sceptic's task, hasn't that defence already been given when Kripke had his sceptic anticipate precisely this move and dismiss it in advance as 'desperate', on the grounds that it left the nature of meaning as mysterious as it was before (K: 51)?

McGinn is not formally committed to endorsing the thesis that meaning is an 'irreducible fact' (M: 151, fn 11), because his stated aim is to set out *possible* responses to the sceptic. Nevertheless, he makes it clear that in his view there may be no way of avoiding resort to such a claim at some point in constructing an account of meaning (M: 175). He clearly treats the appeal to the 'irreducible' facts of meaning as perfectly acceptable, rather than profoundly suspicious or downright disreputable. This preparedness to invoke irreducible semantic categories, rather than explain them in more basic terms, comes out clearly when we see how McGinn believes Wittgenstein himself would have replied to the sceptical challenge: 'What Wittgenstein is saying is that *certain sorts* of facts fail to determine meaning, *viz.* substituting one sign for another, not that *no* facts do (M: 69).

The reference to substituting one sign for another picks up the theme of 'interpretation', and the Wittgensteinian claim that interpretation leaves meaning hanging in the air. In McGinn's words, the lesson to be learned from the sceptical argument 'is not that meaning does not consist in individualistic facts' (M: 69). That is to say: the philosopher who wishes to do so can retain the idea that meaning *is* an individualistic fact. All that need be insisted on is that meaning is an individualistic fact *other* than one of those considered by Kripke, i.e. past practice, images, feelings, dispositions and

the like. This can be done by saying it is a special sort of fact in its own right. For McGinn's individualistic Wittgenstein, meaning certainly *does* consist in an individualistic fact. Thus: 'Wittgenstein *does* suggest that understanding consists in a fact, the fact of having an ability to use signs' (M: 71).

The crucial fact, then, is an *ability* possessed by the individual sign-user. The idea of an ability (such as the ability to mean addition when we say 'plus') is a close relation to that of a *capacity*. The words 'ability' and 'capacity' are almost synonymous, and are treated as such in McGinn's book. We display our ability to use the sign meaning addition by bringing into play our capacity to add. If we then want to know what makes the display a display of adding ability rather than quadding ability – to use the sceptic's example – the answer is: because it is the adding capacity that is being brought into play, not the quadding capacity. We are told:

> it is the concept of *addition* that I exercise when I do computations involving '+', and not the concept of *quaddition*, because the capacity that gets brought to bear is the capacity *to add* and not *to quadd*, where the former capacity is conceived as a capacity to recognise what is the *sum* of pairs of numbers (M: 169–170).

There is much more in a similar vein. For example, the capacity to recognise a number as the sum of two numbers is likened to the ability to see that certain objects fall under a concept. We are then told that the correct use of the concept is 'built into the specifications of the ability' (M: 32 fn 33). Given that McGinn is 'taking capacity concepts as primitive' (M: 174), it is clear that he has not travelled far from the irreducibility thesis. The individualist is telling us that a state of the individual psyche, an act of meaning, of its own nature embodies rule-like properties, in particular, normative constraints. 'If it be asked how this normativity works', says McGinn, 'then the answer (according to the irreducibility thesis) is that it is simply in the nature of meaning to have normative consequences' (M: 163 fn 31).

Faced with these appeals to irreducibility our sympathies should be with the sceptic. Invoking irreducible acts of meaning with which to answer enquiries about the nature of meaning gets us nowhere. Some starting point has to be taken for granted in every explanation, but informative premises will be at some distance from the phenomenon to be explained. If the problem is to understand X,

then an explanation postulating X will be no help. Bertrand Russell said all that needs to be said about the method of postulation when it is used in this fashion. It has, he observed, many advantages. Unfortunately, 'they are the same as the advantages of theft over honest toil' (Russell 1919: 71).

After having canvassed the virtues of the irreducibility thesis, McGinn's argument changes course. He develops his commitment to individualism in a way designed to forge an alliance between individualism and naturalism. The key idea here is that of our natural propensity to act in certain ways. I shall now explore this theme. For McGinn, Wittgenstein is an epistemological naturalist (M: 40), and 'semantic discourse' can be illuminated in terms of what is alleged to be Wittgenstein's 'fundamental thesis' namely, that 'meaning rests ultimately up on the bedrock of our natural propensities' (M: 138). Given the sceptic's challenge and the parts of it that were acknowledged to be successful, it is clear that these propensities are not propensities to form inner states of consciousness, nor are they propensities to replace one symbol by another (i.e. to engage in interpretation). Nor do they consist in intellectual acts of ratiocination. Rather, they express themselves directly in action. In the case of rule following, says McGinn: 'our natural ways of acting with signs provide an adequate basis for the epistemic claims we make about the correctness of what we do; they give us the *right* to have the confidence we actually have (M: 25–26).

The striking feature of this passage, and the feature that should immediately give rise to doubts, is the claim, made on Wittgenstein's behalf, that our natural way of acting with signs provides an adequate basis for normative claims. Thus we are told: 'The basis of the normative is the natural' (M: 86).

What does 'natural' mean here? How broadly is the term to be construed? Does it include not only, say, biological and physiological facts about rule followers, but also psychological and perhaps even sociological facts? Given that, for McGinn, Wittgenstein was *not* 'some kind of "conventionalist" ' (M: 86, fn 32), and that the aim is to drive a wedge between Wittgenstein and sociological answers to the sceptic, we can be sure that 'natural' isn't meant to include sociological facts. This is borne out by the details of the text and the precise ways McGinn uses the word 'natural' throughout his criticism of the sociological approach. McGinn accuses the supporters of the sociological approach of tacitly accepting what Wittgenstein emphatically rejected, namely, that we follow rules by

a process of interpretation. We learn from Wittgenstein that interpretational intermediaries must come to an end, and yet, says McGinn, the sociological approach insinuates society into the story precisely in the role of such an intermediary. It therefore represents exactly the kind of theory Wittgenstein opposed. The issues can be represented diagrammatically. Wittgenstein opposed any model of the fundamental process of rule following constructed along the lines of Figure 7.1. According to McGinn, what the sociologist offers is shown in Figure 7.2.

Any interpretive intermediary, says McGinn, creates a 'gap' between a rule and its application. But, the argument continues, Wittgenstein showed us that, in the last analysis, we must be able to follow at least some rules without interpreting them. Ultimately we must be able to act blindly, and respond to the rule directly. The correct view, therefore, is one in which rule following is based on our natural, spontaneous responses, or on what McGinn calls the 'primitive and instinctual parts of our nature' (M: 120). To represent McGinn's position at this point we need to use another diagram with a different layout. We could either represent the facts of 'our nature' as *underlying* both rule and application, as in Figure 7.3(a), or as *encompassing* them both, as in Figure 7.3(b). Either metaphor would do the job.

McGinn's argument is that, because of the shortcomings of the sociological view, we are driven back to a theory emphasising our individual, natural propensities.

There are two questions needing to be addressed. First, is this characterisation of the sociological view, in particular its handling of interpretive processes, an accurate one? Second, is McGinn's form of individualism, putting all the explanatory weight on our natural propensities, a tenable position? I will argue, contrary to

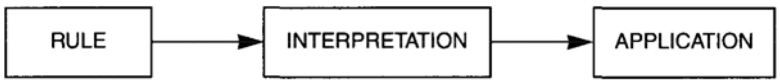

Figure 7.1 The interpretation model

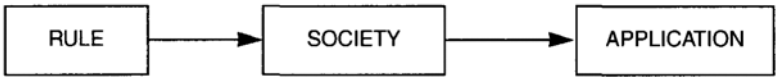

Figure 7.2 The sociological model, as seen by McGinn

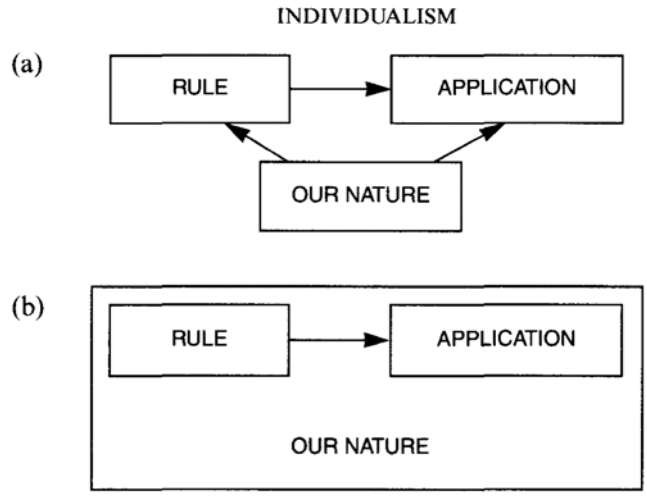

Figure 7.3 The 'naturalistic' view

McGinn, that both questions require a negative answer. My conclusion will be that he has mis-characterised the sociological view, and that the individualistic form of naturalism he offers as an alternative is as untenable as it is un-Wittgensteinian.

Why does McGinn think that a sociological account treats society as an interpretive intermediary, hence reproducing the errors Wittgenstein opposed? The answer is: because he assumes the sociological account of rules always requires the rule follower to *consult* society, i.e. to confer with other people, before making the move between one application and the next. The sociologist is said to picture the rule follower following a given rule by seeking the guidance of another rule, namely: see what everybody else is doing at this point, and then do that. On the sociological view, he thinks, we must always be looking over our shoulders. Thus:

> the community view allows me to get beyond, or beneath, my natural sign-using propensities to something that can be cited to give these propensities a justification . . . the community, in short, provides the kind of *guidance* that Wittgenstein explicitly says there isn't (M: 83).

That guidance, beneath or behind an act of rule following, is allegedly provided, on the community view, by the maxim McGinn characterises as: 'check that this reaction is right by looking to the community' (M: 83, fn 28).

The error of this, McGinn argues, is not only that it opens up the infinite regress of requiring a rule for following a rule: it also fails because it stops us solving the problem of normativity. It means we cannot say what it is to get a rule right as distinct from wrong. This approach only delays the problem or introduces a detour. The rule 'do what everybody else does' only gives us the *right* answer (say, to an addition problem) if everybody else gives the right answer. This rule would give us the wrong answer if the 'community' we followed was itself making a mistake. The real criterion of rightness or wrongness, the argument continues, isn't alignment with others: it is giving the number that really *is* the sum of the other two (M: 185). Given the failure of the sociological view – McGinn concludes – we must go back to the natural inclinations that underlie our use of rules: 'Understanding, we might say, is an unmediated propensity to act' (M: 43; see Figure 7.3).

The misrepresentation involved in McGinn's account of the sociological view becomes apparent once we realise it rests on a conflation. He conflates an *unmediated* propensity with an *unsocialised* propensity. The word 'unmediated' hovers between 'unmediated by processes of interpretation' and 'unmediated by processes of socialisation' (for perhaps the clearest exhibition of this ambiguity, see M: 42 and 43). The assumption is that a socialised propensity is a 'mediated' propensity, and that to achieve an *un*mediated propensity we must keep society out of the story, and go back to nature. This overlooks the fact that a response can be both socialised and, in the relevant sense, unmediated. A response can be the product of training and social interaction, and yet have all the requisite qualities of immediacy, automaticity and unreflectiveness Wittgenstein rightly required, and which he summed up by saying that, in the last analysis, we must follow a rule blindly (PI: 219). It is a commonplace to say of those who have been intensely trained, that what they have been trained to do becomes 'second nature'. Wittgenstein gave an example:

> When someone, whom we fear to disobey, orders us to follow the rule . . . which we understand, we shall write down number after number without any hesitation. And that is a typical kind of reaction to a rule (RFM VI: 47).

Another example he gave was shouting Help! when drowning. We just react in this way: it is instinctive, but it is an instinct shaped by exposure to a particular language. It is at once instinctive and

socialised. McGinn has forgotten our 'second nature', and so thinks – contrary to Wittgenstein – that automatic, unreflective and spontaneous responses must flow simply from our (first) nature (cf. PI: 441). There is, however, nothing in the sociological model carrying the implication that individuals cannot just *act*, or can only act *via* or *after* looking over their shoulders to see what everyone else is doing.

Having denied himself the resources of the sociological approach, what is McGinn's own account of the basis of right and wrong in rule following? What is his analysis of normativity, and his account of Wittgenstein's understanding of it? Here is a representative quotation:

> Wittgenstein sees no real substance to the question whether what we are *naturally* inclined to do *really* conforms with the meaning of our signs: such a question must be futile, since what we are by nature inclined to do is what it is that *constitutes* what we mean (M: 85–86).

We must remember that McGinn intends this passage to be read individualistically, not sociologically. When he falls back on what he calls 'our natural sense of what is right' (M: 88), he means our individual sense of what is right. He has in mind the philosophical equivalent of putting your hand on your heart and saying: we all know what is right, because we feel it here. It was this naive subjectivism that Wittgenstein rejected when he insisted that a standard must be something external and independent (PI: 265). McGinn's individualistic naturalism has turned Wittgenstein into the very opposite of what he really was.

The tendency of the individualist reading to stand Wittgenstein on his head does not stop here. Recall Wittgenstein's insistence that a rule cannot be obeyed just once. If the individualists are right, Wittgenstein must be wrong. If to mean something is to exercise a capacity, then you can surely exercise a capacity just once. The same conclusion follows if meaning is an unmediated natural propensity. We can have a propensity we exercise just once – or even that we never exercise at all. Wittgenstein must therefore be wrong in his commitment to what McGinn calls the multiple application thesis. And this is exactly how McGinn presents the matter. In drawing this conclusion he does not doubt his interpretation of Wittgenstein's individualism: he doubts Wittgenstein's ability to sustain his own (alleged) individualism in a consistent way.

Before reaching this conclusion, however, McGinn discusses a possible alternative explanation of why Wittgenstein adopted the multiple application thesis, i.e. the thesis that a rule cannot be followed just once. Could it be, McGinn asks, that Wittgenstein adhered to a theory of meaning – called by McGinn the 'creative thesis' – and this thesis entails the multiple application requirement? As we shall see, McGinn's 'creative thesis' is recognisable as a version of what I have earlier referred to as 'meaning finitism'. McGinn, however, rejects the creative thesis, and hence meaning finitism, as wrong and as inconsistent with Wittgenstein's position. Clearly, we must examine his reasons.

The 'creative thesis', just like finitism, implies that meaning is created in a step-by-step way through acts of use. Meaning does not exist independently of use or in advance of it, but is created as we go along. This explains why multiple applications are needed. Without it nothing has been created, and so meaning simply would not exist. Unfortunately, McGinn does not confront meaning finitism in its general form, but insists on discussing a special case. To make matters worse, the special case in question is not well suited to capture the general character of Wittgenstein's thought. He defines the creative thesis by analogy with 'existentialist' ideas drawn from Sartre, thus making it out to be, in his own words, a 'somewhat obscure and heady claim' (M: 134). The creation of meaning is likened to the creation of our individual characters and personalities through free acts of self-definition, or existential choices. Thus McGinn tells us:

> The idea is reminiscent of the existentialist conception of a person: it is the free action of an agent that *constitutes* what he is; it is not that action springs from an antecedently constituted self. . . . The 'existentialist conception of meaning' is thus the thesis that linguistic behaviour is free and unconstrained, and that we can speak of determinate meaning only as the *upshot* of such behaviour (M: 134, fn 56).

The error here is that there is nothing about the general form of the creative thesis, i.e. finitism, making it a celebration of 'freedom' or lack of constraint. In so far as meaning deserves the epithet 'indeterminate', this could be generated by its dependence on an endlessly unfolding sequence of contingencies. According to the creative thesis as McGinn defines it, 'the meaning of a word is thus in some way indeterminate until the sum-total of its use has been

reckoned with' (M: 134). This cannot be an apt way of expressing things, because there is no such sum total. (It is like talking about the 'final price'.) There is always, and only, the running total so far. We should *either* say that meaning is always indeterminate, *or* that the correct notion of what it is for meaning to be 'determinate' is pragmatic and contextual. For example, we might call meaning 'determinate' in so far as it functions acceptably within some practical context. Meaning would be 'determinate' when it is determinate enough for immediate, practical purposes.

McGinn dismisses the idea that Wittgenstein subscribed to the creative thesis as no more than an 'ingenious' suggestion. The reason given is that the creative thesis cannot explain the normative character of meaning and rules. As McGinn sees it, on the creative thesis, what I mean at any given moment is not 'determinate'; therefore 'what I now mean cannot normatively constrain how I subsequently use the word in question' (M: 134). But, alleges McGinn, Wittgenstein rightly and explicitly insisted that 'future use *is* determined by present meaning' (M: 134). Hence, he concludes, Wittgenstein didn't hold the creative thesis, and so this cannot have been his reason for saying that we cannot obey a rule just once. His commitment to the multiple application thesis must have been no more than an aberration.

Did Wittgenstein really say future use is determined by present meaning? Isn't his position that use determines meaning, not that meaning determines use? McGinn's evidence for his inversion of the usual reading is a well known passage discussing the idea that usage derives from our grasping a 'sense'. Someone possessed by this picture of language is represented in the passage as saying subsequent usage is, in a queer way, already present in an act of meaning. Wittgenstein replied: 'But of course it is, "in *some* sense"! Really the only thing wrong with what you say is the expression "in a queer way". The rest is all right . . . '(PI: 195).

In this passage we appear to have been told usage *is* already present – at least in some sense – and, provided we don't think of its being present in a queer way, it is entirely permissible to say it is already present. This, McGinn argues, constitutes a denial of the creative thesis, because on the creative thesis usage isn't already present, and cannot be said to be present even in a non-queer way.

This is weak evidence, and in citing this passage and seeking to use it as he does, McGinn is clutching at straws. His gloss on PI: 195 was directly contradicted by Wittgenstein in PI: 197 (and repeated

in RFM I: 130), which makes explicit the grounds for the usual reading. Rather than saying there is a genuine sense in which usage can be present, i.e. present in a non-queer way, there Wittgenstein clearly said it was the very idea of usage being present at all that constitutes the objectionable queerness. Discussing the phenomenon we call grasping the use of a word 'in a flash' he said, 'It becomes queer when we are led to think that the future development must in some way already be present in the act of grasping the use and yet isn't present' (PI: 197; RFM I: 130). Notice, Wittgenstein didn't say it becomes queer when we are led to think of it being present in certain incorrect and suspect ways, but is non-queer when we think of the presence in other ways. He said it becomes queer when we are led to think of it as present in some – i.e. any – way (*auf irgendeine Weise*). The most that could be said for McGinn's reading is that he is exploiting an apparent contradiction between two passages in Wittgenstein. Even on this account of the situation, he has come down on the wrong side. The passage on which his argument depends is only one of a number whose overall tendency is quite clear, and which work against his interpretation (cf. RFM I: 126; Z: 173; PI: 334). They remind us that, for Wittgenstein, future use does not really and literally flow out of the process of 'grasping a sense'. The future applications of a term are not 'in some *unique* way predetermined, anticipated – as only the act of meaning can anticipate reality' (PI: 188). This is the mythological picture of meaning determinism, the one Wittgenstein warned us against (RFM I: 119).[1]

Even the appearance of a contradiction between PI: 195, the passage quoted by McGinn, and the overall thrust of Wittgenstein's argument, can be quickly dispelled. It arose because Wittgenstein seemed to be making a concession to the idea of usage being present in the act of meaning. He said: of course it could be said to be present in *some* sense. This can be explained by his not wanting to interfere with ordinary usage, or to bully us into changing the way we talk – provided that talk is of an unselfconscious kind taking place in an everyday context. That he had in mind a mere manner of speaking, rather than something to be taken literally, is made clear by the rest of the passage in which he instanced a child's misunderstanding of a familiar phrase. The same message is driven home elsewhere. The words 'You already know how it is', Wittgenstein indicated, is just what we say, just the form of words we use, to gloss the unhesitating character of some of our rule

following (RFM VI: 47). We are inclined to say things of this kind, and there is nothing wrong with doing so, provided we don't mistake these natural metaphors and figures of speech for genuine insights or philosophical truths. Thus:

> Of course, we say: 'all this is involved in the concept itself', of the rule for example – but what that means is that we incline to *these* determinations of the concept. For what have we in our heads, which of itself contains all these determinations (RFM VII: 42)?

The Wittgensteinian answer, contrary to McGinn, for whom future use is determined by present meaning, is that nothing in our heads could possibly contain all these determinations. This is not to say, when we are given the order 'add 2' to produce the sequence of even numbers, we are in any doubt about what to do after, say, 2004. We answer 2006 without hesitation. But, to repeat a passage cited earlier, this does not mean it is answered in advance (RFM I: 3). Furthermore, Wittgenstein's position permits a solution to the problem of normativity which is perfectly compatible with meaning finitism, or what McGinn calls the creative thesis. Constraints can be generated as and when they are needed, without the superfluous assumption of pre-existence. Our confident behaviour, when taken collectively, is sufficient to generate the requisite consensus on a step-by-step, case-by-case basis. This can provide an independent standard – independent, that is, from any individual whose confident use of the rule may turn out to be misguided.

Once again McGinn's individualistic reading has turned Wittgenstein on his head. He has made Wittgenstein out to be committed to the doctrine identified earlier as meaning determinism. This is the opposite of everything the later Wittgenstein struggled so hard to establish and make plausible.

8

ISOLATION AND INNOVATION

The individualist analysis of rule following is sustained by the intuition that a person who is physically isolated can still follow a rule. This alone seems to prove there must be something wrong with any collectivist account. As McGinn puts it:

> What we need for a genuine community conception is the idea that following a rule is a notion, like that of being in fashion or marching in step, that inherently involves reference to individuals other than the individual to whom the ascription is made (M: 194).

Robinson Crusoe, alone on his desert island, cannot march in step or out of step, because that implies co-ordination, or lack of it, with fellow marchers of whom, by definition, there are none. Nor can he be a follower of fashion, or indifferent to fashion. The choice of fashionable or unfashionable clothes is made by reference to what other people are wearing, and there are no such persons on his island to provide a reference group. As soon as Crusoe finds himself isolated the possibility of his marching in step disappears. It immediately evaporates, as does the possibility of his being fashionable, and yet we don't think his words suddenly lose their meaning or his ability to do arithmetic evaporates in the same way. That rule-following abilities linger on in isolation suggests they are qualitatively different to other activities like marching in step or following fashion. If the latter are, to use McGinn's terminology, 'essentially social', and disappear in isolation, doesn't this suggest rule following can't be essentially social?

A further example of an essentially social activity is commerce. By definition commerce involves a buyer and a seller and Crusoe cannot be the sole buyer and the sole seller – not, at least, if we are

using words in the ordinary way. He can no more genuinely buy things from, and sell things to, himself than his right hand can give his left hand a gift. He cannot engage in commerce with himself. Wittgenstein asked whether the question 'Could there be only one [human being] that followed a rule?' is like 'Can one man alone engage in commerce?' (RFM VI: 45). Unfortunately, he did not write down his answer. An individualist would reply that the questions may be of the same form, but the answers are different: following a rule is not like engaging in commerce because commerce is essentially social and rule following isn't. In contrast to this, I am committed to saying the answers to the two questions are similar. Following a rule is, in vital respects, like engaging in commerce.

One error in the individualist position derives from the failure to draw an important distinction: physical isolation is a different phenomenon from social isolation. Total physical isolation, from birth to death, would entail social isolation, but physical isolation, as such, does not entail social isolation. These two things are of a quite different nature. The difference is made evident by the fact that social interactions are typically episodic, containing gaps between periods of face-to-face encounter. For the duration of these gaps we are, or may be, physically alone. Physical isolation does not therefore necessarily destroy, terminate, or preclude a continuing social interaction. Face-to-face interaction is, presumably, the primitive scenario in which institutional structures and conventions are first created. But even in face-to-face interactions the vital ingredient in the process is not of a physical but of a cognitive character. Institutions and conventions arise because of all the primitive and instinctive predicting, calculating, knowing, remembering and referring that takes place in the course of interaction. The cognitive dimension, as it may be called, must be given due recognition. The social character of a situation arises when we have interactions that are informed by expectations and some measure of shared understanding. Indeed, for the sociologist, these components enter into the definition of what it is for an encounter to be an 'interaction' rather than mere physical proximity.

Driving on the left-hand side of the road is the convention in Scotland. The convention is sanctioned by law, but that is a complication we can leave aside. A convention is an essentially social activity, being a regularity in behaviour predicated on the condition that others conform to it as well. If, as a Scottish driver, you came to believe that other drivers were going to drive on the right-hand

side, then you would do so too. Following, rather than merely incidentally conforming to the convention, thus involves an essential reference to others. Now, driving appropriately along the left side of a characteristically deserted Highland road still counts as adhering to the convention. How could this be if isolation as such were destructive of an essentially social activity? The answer, clearly, is that not all social activities evaporate given physical isolation. This is because of the different degree and manner in which cognitive processes such as memory, belief, and expectation are involved in different activities. This varies widely in different situations. As a consequence the conventions governing the frequency and intensity of the face-to-face component of an activity are very diverse. Marching in step is a case where the requirements concerning spatial contiguity are especially restrictive, which is why it is unhelpful to take it as a representative model of social processes. If this case is uppermost in our mind we might mistakenly think any increase in spatial separation would necessarily attenuate the social element in a situation. In general, however, interactions are not tightly constrained by their spatio-temporal parameters. The frequency of interaction, the interval between face-to-face meetings, the length of time before expectations come into play or are put to the test, are highly elastic and highly variable between institutions.

Consider Wittgenstein's example of commerce. A physically isolated person such as Crusoe cannot engage in commerce *with* himself, but he can engage in commerce *by* himself, just as the Scottish driver can conform to the rules of the road by himself. An isolate can be engaged in commercial projects aimed at distant and not yet identified others, as are explorers, prospectors or isolated hill farmers. This is well brought out in a poignant passage in Defoe's story. Rummaging around the wrecked ship to see what he can salvage, Crusoe discovers a locker with some razors, a pair of scissors and some knives and forks in it – and 'about thirty-six pounds value in money', some of it in gold coin, some silver. He eagerly takes the useful implements but smiles to himself on seeing the money. What good is this to him? 'I have no manner of use for thee; even remain where thou art, and go to the bottom as a creature whose life is not worth saving.' But Defoe follows this up with a little twist. He makes Crusoe say, 'However, upon second thoughts, I took it away' (Defoe 1719: 70). The point is not explained, but every reader will understand what underlies these second thoughts. Crusoe was planning for a possible, though perhaps unlikely, future

contingency. Action in the present is mediated by thoughts about the past and the future. Using his retrospective knowledge he behaved in his isolation in the light of prospective contingencies and imagined future interactions. This was the action of someone whose isolation was physically real, but whose society still entered into his calculations and structured his thoughts and actions. It is surely reasonable to call Crusoe's action in taking the money 'essentially social'. It inherently involves a reference to (unknown and unspecified) individuals other than the one to whom the action is imputed. The action is the expression of a thought, the thought can only be explained by its content, and that content only by its reference, and the reference is to a future commercial transaction. It would be better to take this action of Crusoe's, despite his physical isolation, as a model of social action in preference to McGinn's special case of marching in step.

Within broad limits, then, a defender of the sociological approach to rules can easily make sense of, and justify, our intuition that physical isolates, like Crusoe, can follow rules. Crusoe's thoughts place him, as it were, in the familiar gap between face-to-face encounters – a gap that can be indefinitely prolonged without becoming a social vacuum. But what of the contrast between the rapid evaporation of the possibility of marching in step, by comparison with the lingering on of the ability to follow rules? Does not this attest to the qualitatively different nature of the two cases, the one being essentially social, the other not? The answer is no. It attests to the difference between the spatio-temporal parameters conventionally applied to two different sorts of essentially social activity. The individualist and the collectivist thus agree that in principle Crusoe can follow rules, but they reach this conclusion by different routes. For the individualist it seems obvious that Crusoe can follow rules because rule following is nothing more than forming an intention and carrying it out, where intentions are taken as things with intrinsic propositional content generated by the individual mind. For the collectivist the powers intrinsic to Crusoe's mind, though necessary ingredients in the process, are not sufficient. Crusoe can follow rules because these intrinsic powers are augmented and set in a social context by extrinsic circumstances, though these extrinsic circumstances can be present even in physical isolation. The basic requirement is that the rule follower be a social agent and a member of a community following the rule in question. For the reasons given, Crusoe can fulfil this requirement.

Individualists will resist this line of argument, suspecting that the claim that rule following is essentially social is becoming trivialised. This is McGinn's complaint against Kripke's version of collectivism. Kripke says the sociological approach presents us with no problem in seeing Crusoe as a rule follower: all that is required is that we look at Crusoe's actions as we would those of a typical and unproblematic member of society. When we think of Crusoe engaging in some action, or uttering some sounds, we react as we would react if those actions and sounds were produced in an ordinary social setting. By thinking about him in this light we are in effect 'taking him into our community', and hence satisfy the collectivist requirement that other people are involved in the imputation of rule following (K: 110). But if we say rule following is only possible for a member of a community, and then make membership *that* easy, are we not voiding collectivism of real content? On the other hand, if membership is construed substantially, rather than notionally, the collectivist position seems clearly false. For example, suppose we formulate the claims of collectivism by laying down the following principle: *before someone can truly be said to follow rule R they must belong – really belong – to a community whose other members also follow R*. This principle has been called the 'strong thesis'. The strong thesis, critics have argued, makes collectivism significant, but at the same time clearly shows it to be false. It allows us to say that Crusoe could follow the rules he had acquired through his previous socialisation (provided we accept that his orientation, and our hypothetical response to it, amounts to continuing membership), but it precludes his doing what he surely can in fact do, namely inventing new rules for himself. Indeed, the strong thesis would prevent any single individual, whether alone like Crusoe or in the midst of an intense communal life, from inventing new rules. It makes it unintelligible how anyone might introduce a new rule into a rule-following community. Innovation involves the innovator going out on a limb by being the first to follow a new rule, but that puts the innovator in violation of the strong thesis. If the thesis is right, innovation is impossible, but innovation is possible, so the thesis is wrong. As McGinn says:

> I take it as obvious that this strong thesis is self-evidently absurd, and I doubt that it has ever been (explicitly) held: it makes nonsense of the idea that a member of a rule-following community can be *innovative* in the rules he follows (M: 195).

This may be called 'the argument from innovation'.[1]

I shall defend the strong thesis and show that, far from being 'self-evidently absurd', it must be the basis for our understanding all cases of rule following, including Crusoe-like cases and innovations. To make good this claim we must cast a critical eye on the argument from innovation. We will see that the strong thesis does not, in general, rule out innovation, it only precludes innovation as the individualist wrongly conceives of it. It is the inadequacy of the individualist analysis of innovation that stands in contradiction to the strong thesis, not the phenomenon itself. McGinn tries to make the argument from innovation plausible by inviting us to consider examples from the sciences. Before turning to these it will be useful to address a simpler and older version of the argument drawing upon everyday categories.

It seems evident that Crusoe could give names to the objects on his island. He could surely name landmarks or unusual kinds of flora and fauna. A. J. Ayer developed this theme when he asserted that Crusoe could see a new kind of bird and give it a name (Ayer 1954). To give an object a name is to create a rule, namely the rule: this object, or kind of object, has the name N. If Crusoe can do this he violates, and hence refutes, the strong thesis that an individual can only follow rule R if he belongs to a community where others follow R as well. No one in Crusoe's past has, or need have, used the names he has just introduced. We may suppose he is a true innovator in these acts of naming.

The choice is stark and simple. Either Crusoe's naming refutes the strong thesis or we must say Crusoe cannot introduce names in the way we are supposing. The collectivist must not duck the issue. Collectivism does indeed have the implication that Crusoe cannot really or strictly be doing what we think he is doing, or what we think he is doing when we are under the influence of the individualist argument. Indeed, it carries the implication that no single individual, whether alone or interacting with others, can achieve such a result by depending entirely on their own resources. The collectivist does not deny that we say Crusoe can innovate and name, or that we speak as though these are individual accomplishments. The claim is that in saying such things we are speaking superficially and uncritically and, for present purposes, wrongly.

Here is how the collectivist argument should go. The important step is to realise that innovation, even the simple innovation of giving something a name, is a process. Being a process it has an

inner structure, and that structure is both historical and social. The process can be divided into two main phases. The first may be called the 'initiation', the second the 'culmination'. Thus someone may propose that an object has such and such a name, or that such and such a procedure become a rule, or that some technique be adopted and understood in a certain way. This act of initiation may go no further: the proposal may fall on deaf ears. Alternatively others may begin to take up the suggestion and model their practices on the new exemplar. Its use may spread and become the accepted currency of interaction. Now the innovation would be complete. This would be the culmination: the innovation would have become a veritable institution. These considerations apply both to names for kinds of things and to proper names. (For a valuable reminder of the social significance of proper names see MacIntyre 1988: Chapter XIX.)

Suppose that, instead of appreciating its complexities, we had an uncritical and simplistic grasp of innovation. Suppose we compress the process, in our understanding, into an event. Instead of seeing its inner structure we treat it as if it were an unstructured point. Perhaps, using Ryle's terminology, we say to ourselves: innovation is a 'success word', it refers to an achievement not a process, it is like *winning* a race, not *running* a race (cf. Ryle 1949: 149). Such a stance would collapse together the phases of initiation and culmination. Everything would seem like a moment of pure culmination. From the standpoint of the process picture such ideas, and the talk they encourage, are shot through with ambiguity. Uncritical references to innovation hover between innovation-as-initiation and innovation-as-culmination, with no clear sense of which phase of the process is meant.

These considerations can now be brought to bear on the argument from innovation. Remember that for the collectivist a name is a social institution: to have a name is to have a social status. If a critic of collectivism has in mind the initiation of the process of naming then, of course, it can in principle be imputed to an individual. Such individuals may have very little sense of what they are setting in motion. They may act routinely, or unselfconsciously, or unwittingly. Often, though, the act of initiation will be done with the hope, expectation or assumption that the initial move will be taken up and accepted by others. Talking as if the process has already reached culmination, even when it hasn't, is one way of helping it on its way. Each individual act helps to put one element in

place in a reality constituted of such acts. If the process does reach culmination, if the individual move becomes a collective practice, and the name an institution, then the initial act will have 'primed' the institution. At this stage in the process, be it naming or some more complex technique, the events clearly cannot be located in, or imputed to, any given individual. This conclusion holds even of the individual, if there is one, who was responsible for setting the process in motion. The individualist analysis, therefore, may make sense for the phase of initiation, but certainly not for the phase of culmination. Crusoe the innovator, like any innovator, was engaging in activity that could in principle initiate or prime the name-as-institution, though nothing he could accomplish alone could itself constitute the institution. By focusing our attention on just one phase of the process of innovation, the individualist tries to give plausibility to a conclusion which, to be tenable, would have to hold for the entirety of the process. It does not so hold. On the contrary, it is the strong thesis that embodies the conditions necessary for the culmination and achievement of innovation.

We have now seen how the individualist case appears to, but does not genuinely, derive strength from the appeal to innovation. It trades on an ambiguity which the individualists are responsible for importing into the discussion. Individualists, of course, see no such ambiguity. For them the initiatory phase is all: that *is* the moment of innovation. In adopting this position they take themselves to be simply reporting facts, and reflecting our pre-theoretical intuitions. For the collectivist the voice of common sense does not speak in such clear and partisan tones. For them our willingness to say 'Crusoe introduced the name N', is on a par with our casual willingness to say, 'I promised myself X'. Though it may be harmless and idiomatic it is strictly wrong and should carry no weight in a serious analysis of the phenomenon. The strong thesis, then, is inconsistent with some uses of the word 'rule'. Logically these must be deemed false or mere metaphors and analogies. Methodologically the important thing is that the strong thesis be based on, and motivated by, an illuminating model of rule following. This condition is satisfied by appeal to the self-referential model of an institution.

Wittgenstein always emphasised the need to select the right focal point for understanding and enquiry. A fruitful 'paradigm', 'object of comparison', 'prototype' or '*Urbild*' was vital (cf. CV: p.14, p.26; PI: 122, 130, 131, 385). It is also clear that, in his view, the prototype for a concept such as following a rule, like the other terms

describing human conduct, is to be extracted from real-life cases central to our everyday activity, not just from thought experiments and our imagined responses to them. This means selecting our prototypes from concepts used for describing collective behaviour or behaviour within an interacting collective. Thus, when Wittgenstein asked himself how human behaviour should be described, he answered,

> Surely only by sketching the actions of a variety of humans, as they are all mixed up together. What determines our judgement, our concepts and reactions, is not what *one* man is doing *now*, an individual action, but the whole hurly-burly of human actions, the background against which we see any action (Z: 567).

In this passage Wittgenstein wasn't explicitly talking about rule following, but about identifying a particular sort of behaviour, such as 'pretending'. Nevertheless the argument is general and applies to rule following. The important thing, he said, was seeing a pattern:

> Seeing life as a weave, this pattern . . . is not always complete and is varied in a multiplicity of ways. But we, in our conceptual world, keep on seeing the same, recurring with variations. That is how our concepts take it. For concepts are not for use on a single occasion (Z: 568).

Wittgenstein's talk of seeing life as a weave is hardly the language of an individualist, asserting, as it does, the priority of the overall pattern of social life over the meaning of its individual components. Each individual episode is understood as being part of an overall weave, with the individual threads of action appearing and disappearing like the warp and weft of a fabric. These passages are reminiscent of von Mises' 'indissoluble intertwinement of actions and reactions' in the 'coherent and indivisible' structure of market relations. Notice also how Wittgenstein characterised the pattern as one requiring for its detection the participation in a 'conceptual world' (*Begriffswelt*). He did not explain what he meant by this idealist sounding phrase. One interpretation is provided by the collectively created, self-referring patterns described in our model of a social institution. Institutions, on that model, could be called 'conceptual worlds'.

Wittgenstein was well aware that our casual talk of rules, as of other things, is not confined to central, unproblematic, or paradigm

cases. We extend labels to cases that do not contain all the features of the prototype (cf. Z: 112; CV: 14). This practice can generate apparent counter-examples to a sociological analysis, as when we speak of a person issuing a rule to *themselves*, and then obeying it. Wittgenstein's response was straightforward and effective: 'Certainly I can give myself a rule and then follow it. But is it not a rule only for this reason, that it is analogous to what is called "rule" in human dealings' (RFM VI: 41)?

In these cases we are not, therefore, dealing with real counter-examples but with mere analogies. By 'human dealings' I take Wittgenstein to be referring to the hurly-burly of life mentioned in the previous quotation. Here the notion of a rule, and the success or failure of particular acts of rule following, are the currency of social interaction. The implication is clear: the way to study the phenomenon of rule following is to look at what is called a 'rule' in all its complexity and richness. The individualist wants us to focus on marginal and simplified cases – mere analogies – and treat these analogies as paradigms. Nothing could be more calculated to lead us astray (cf. RFM VI: 48; PI: 142).

What would Crusoe's, or anyone else's, behaviour look like if we did not set it against the background of paradigmatic social interaction? Suppose we deliberately didn't mix it up with the hurly-burly of life or weave it into the pattern of our institutions and conventions. What would we be left with? For the individualist, such an exercise would make little fundamental difference. We should still be left with rules and rule following because these are properties of individuals, not properties of the institutions in which they participate. For the collectivist, however, such an exercise would produce a profound qualitative difference. Rules would disappear from view, and all we would be left with would be outer regularities, inner dispositions and subjective experiences. And perhaps once again we would find ourselves in collision with our intuitions. Putting ourselves in Crusoe's place we might be convinced that we could still name things. To use Ayer's example, I could imagine myself looking at the strange bird, concentrating on it, and undertaking to call it by such and such a name. After all, I can see the bird, and that doesn't involve other people. If I can register it, remember it, and match what I see next time to my memory, what more is involved in naming than attaching a label to the memory?

However tempting an account along these lines may seem, a vital part of the story is missing. What has been described is the psycho-

logical machinery by virtue of which the individual is able to participate in the collective process of using a name. The missing element in the story is precisely the collective machinery which allows us to speak of the name being used rightly or wrongly. Ayer's example takes us back to the features of Wittgenstein's position outlined in the chapter on finitism, namely, the problematic move from case to case. We are tempted to say that if the next sighting fits or matches our previous experience, then we have seen the same thing; if it doesn't fit, then we have seen something different. The problem is this fitting and matching. Wittgenstein took the example of recognising a cube. Suppose we match the seen cube with a picture that comes before the mind. In what sense, he asked, can this picture fit, or fail to fit, how we use the word 'cube'? If the answer seems simple, perhaps we have forgotten how many different ways the picture could be interpreted (PI: 139). But doesn't our past experience naturally generate expectations? Surely there can be a clash between a picture in our minds and something we see? Crusoe, like anybody else, and with a similar outcome, could automatically register a sense of difference without going through any interpretive process. Wittgenstein concedes all this: 'people in general apply *this* picture like *this*'. We have here, he said, a normal case, and abnormal cases (PI: 141). But whatever the truth about the normal human reaction, we have not yet provided any structure for distinguishing right from wrong recognitions. We have a subjective distinction but not yet an objective one. No doubt such a structure will be built upon, and built out of, what is normal rather than abnormal, but a statistical concept of normality is not to be equated with the evaluative concept of normativity. Normativity requires a shared language-game, hence, 'a thing . . . has not even *got* a name except in the language-game' (PI: 49).

The danger awaiting the collectivist at this point is over-reaction. If individualists make their case by inflating the subjective, psychological side, and ignoring the sociological side of the story, this invites a response with the same faults in reverse. I have argued elsewhere that the debate over Wittgenstein's account of rules has been dogged from the outset by over-polarisation (Bloor 1989). Collectivists must not gratuitously minimise our innate capacity to act in the world as if, without society, the human mind would be in a state of confusion. Order and coherence, as such, do not arrive on the scene only with the construction of collective representations. (Though it is true that a new dimension of cognitive order then

becomes available, characterised by its impersonality, objectivity, externality and necessity.) Given the impressive competences of the non-human animal mind the confusion hypothesis is implausible. Fortunately, it is not a necessary feature of the collectivist position. The point is not the degree of organisation innate to individual behaviour, but its status. Crusoe could presumably build himself a house, but no matter how sophisticated its structure it would not have the status of a piece of property – and making it yet more sophisticated would not take it one jot nearer to that status. Similarly, as a normal human specimen, he might naturally make all manner of subtly structured and psychologically useful responses to the objects around him. He will surely utilise the psychological apparatus that we normally utilise when we use names, but what is at issue is whether those responses would thereby have the status of names.

How significant is this question of status? If society simply legitimates what can, and must, be done for other (psychological) reasons, if it is just a coat of paint on an otherwise unchanged object, then it isn't a very interesting phenomenon. The individualist could develop such an argument by following Bennett (1961), Stroud (1965) and Will (1985). Their claim is that we don't have any real options in how we follow the basic rules of logic and arithmetic, so society's endorsement doesn't alter anything. The way our brains work, and the way reality is, are sufficient to force us to certain conclusions which no mere conventions could modify or challenge. We may not be able to provide ultimate justifications for the way we infer, or add, but there are no live alternatives: 'he may indeed *say* it, still he can't *think* it' (RFM I: 116). To do otherwise than we do, or try to do otherwise, would be mere clowning, or a recipe for chaos. We cannot seriously entertain the idea of rational beings with an entirely different mentality to ours, finding trivial what we find vastly difficult, whilst not seeing what is obvious to us. Such radical deviance is inconceivable, a *dead* possibility.

There is much truth in this. Wittgenstein would surely have gone along with at least some of it. He did not believe that thinking or inference or counting was open to arbitrary definition or redefinition. Its role in our life was too central for this (RFM I: 116). I interpret this to mean that our innate propensities (to infer and extrapolate and respond to pattern) give certain cognitive structures a natural salience. For good naturalistic reasons this salience will be reflected in the conventions that emerge as individuals interact and

seek to co-ordinate their behaviour. (In Schelling's (1960: 73) phrase they provide solutions to co-ordination problems that nobody can keep their eyes off.) No doubt this is why Kripke's quus function, as we encountered it, is parasitical on the plus function, being defined as a variant of addition. On inductive grounds we would otherwise anticipate insuperable problems getting people to co-ordinate themselves around the discontinuity in such a function, or what appears as a discontinuity from our point of view. There is no special problem once we can communicate using the plus function as background knowledge, but it is difficult to see how things could be done the other way round (cf. Hesse 1974: Chapter 3).

To acknowledge the force of these points, however, is not to concede the individualist objection. Stripped of its naturalistic trappings the objection of Bennett, Stroud and Will is just a variant of the old appeal to self-evidence. The truths it is defending are represented as compellingly evident because their negation is unthinkable. Once seen in this light the reply is clear. What we can, or cannot, conceive is an historical product not a timeless universal. There is no hard and fast line to be drawn between the thinkable and the unthinkable, just as there is no hard and fast line to be drawn 'between what is still and what is no longer called "regularity" ' (RFM I: 116). Our innate limitations may determine our starting point, but they do not tell us where our thoughts may finish up, or how our cultural institutions may be extended and elaborated. Cultural elaboration may take us a great distance from our starting point, and even contradict its indications. A child's innate sense of regularity does not prevent later, culturally refined perceptions of a vastly different kind. Suppose, for the sake of argument, that our innate psychological tendencies made it natural for cultures to develop spatial ideas with a core of (roughly) Euclidean concepts, or dynamical ideas of a recognisable Aristotelian character. This could be true without precluding a step-by-step historical development culminating in the achievements of a Riemann or an Einstein, achievements which would be meaningless in terms of the original ideas. There have been too many appeals of this kind to the naturalness and self-evidence of current cultural forms, and the 'inconceivability' of alternatives, for this manner of argument to deserve credibility.

What of the more specific point about normativity? Why cannot a psychological naturalism explain the normative side of our representations and rules? What are our brains for, if not for adapting us

to our environment and ensuring that our behaviour conforms to the demands of reality? Don't these ideas in themselves carry sufficient overtones of normativity? Perhaps the most determined, and plausible, attempt to date to give a positive answer to these questions is provided by Ruth Garrett Millikan (1984 and 1990). Millikan uses a biological, evolutionary and functional approach to try to accomplish what McGinn aimed for: a way of grounding the normative in the natural, where 'natural' means our biological nature but not our social nature. In thinking about an organ, the heart for example, it becomes clear that functional concepts imply a distinction between what an organ does (in fact) and what it ought to do (according to its functional definition). Here, perhaps, we have a naturalistic basis for our talk of 'ought' and 'ought not'. A heart, argues Millikan, isn't just a thing with certain causal powers and dispositions; it is a thing with a certain function to perform. It has this function, not in virtue of what it *does*, but in virtue of what it *ought* to do, i.e. in virtue of what it is *for*. And what it is for is defined by its evolutionary history. A heart is something that has evolved for pumping blood, and this is still true even if the object in question doesn't and can't pump blood. (A diseased or dead heart is still a heart.) Millikan then applies these ideas to rule following. She suggests that while sophisticated human rule following may be mediated by interpretations and representations of the rule, in the end, it must devolve down to a range of blind or automatic actions. These are innate rule-following abilities we share with lower organisms, right down to humble insects such as the hover-fly. (Biologists, Millikan explains, have identified the precise rule according to which male hover-flies select their trajectories when aiming to intercept female hover-flies.) Our most basic act of rule following, just like the hover-fly's basic rule following, has a right and a wrong attached to it because it is the product of mechanisms which have a proper biological function. These functions are said to receive their specification through the evolutionary processes responsible (in more or less roundabout ways) for the proliferation and presence of the mechanisms in surviving organisms.

Millikan's account is ingenious, and certainly fills in details McGinn's appeal to nature left blank.[2] Will it do as a reply to Kripke's sceptic? Does it, as Millikan thinks, provide a straight, biological answer to the sceptic's challenge by locating an individualistic fact of meaning? The answer is: no. This line of argument does not really allow us to avoid the appeal to society as the basis

for normativity. The reason is very simple. The idea of getting a piece of rule following right, or the idea of getting it wrong in particular cases, are *actors' categories*. They are ideas invoked by rule followers, or would-be rule followers, about one another or themselves. Normative ideas cannot be external to the thoughts and actions of a community of rule followers in the way a fact about its historical evolution might be. The actor's own awareness of these norms is constitutive of their very existence as norms, in the way described in Chapter 3. This is the difference between a sociological fact and a biological fact. A fact of biological reality does not depend on being known in order to exist. A sociological reality, such as an institution, is so dependent. To put the point in another way, a biological fact about evolutionary function is no use as a norm or standard unless it is known and believed and thought of *as* a norm or standard. But then it ceases to be just a biological fact, and assumes a sociological dimension. It becomes part of the culture of those who know it, and only by means of this transformation could it play a genuinely normative role (cf. Haugeland 1990: 414–415). If we are to use functional ideas at all it must be sociological, not biological, functions that are paramount. This is how, on occasion, Wittgenstein expressed himself, saying: 'Compare the meaning of a word with the "function" of an official. And "different meanings" with "different functions" ' (OC: 64; cf. OC: 655).

Official, administrative 'functions' are the tasks an official is supposed to perform and is authorised to perform. These definitions of role and empowerment are sustained as social institutions in ways needing to be understood in terms of the self-referring and performative model.[3]

McGinn's version of the argument from innovation is not based on naming but points us towards scientific innovations. We are invited to consider a 'creative mathematician' discovering and investigating a new function, and a zoologist who 'comes across a hitherto unknown species and gives it a name' (M: 195). McGinn does not analyse these examples in any detail. Perhaps the force of the argument is taken to lie in obvious and general features of the discovery process so that details don't matter. Whatever the reason, he proceeds as if we need only be reminded of such cases to see the falsity of the strong thesis. The previous argument shows that such cases will only carry this message for those who have no conception of innovation as a social process. Investigation in the history and sociology of science over the last few decades has clearly illustrated

the principle that, while an individual may initiate a scientific discovery, the process is never confined to the psyche of the innovative individual. It is widely acknowledged that a new idea can rarely be traced back to a single originator, but that is not the issue here. The important point is that the status of a discovery claim is not settled by the opinions and judgements of a single innovator. It is only settled when the scientific community has established a consensus. McGinn's talk of a zoologist 'coming across' a new species and 'giving it a name' signally fails to do justice to the processes involved. It blurs over the process character of the discovery and encourages us to treat it as a point event. To begin with, species are not things we can come across, though we might come across a finite number of members of a (putative) species. Fortunately no historian of biology would talk in this way.[4] The same applies to his 'creative mathematician' as detailed studies of, say, the reception of Zermelo's axiom of choice, or Cantor's transfinite arithmetic make clear, see Dauben (1979) and Moore (1982). The point is not that all mathematical innovations are like these in causing controversy, they certainly are not. As Wittgenstein pointed out, the process is typically one of 'peaceful agreement' (RFM VI: 21). Nevertheless, all cases share with these the need to distinguish between individual opinion and personal conviction on the one hand, and on the other, the collective endorsement of the proposals. In many cases the transition from the one status to the other is smooth and unproblematic, but that does not destroy the theoretical or practical importance of the distinction.[5]

Thomas Kuhn is the obvious representative for this approach in the history of science. As he put it:

> It is not, after all, the individual who decides whether his discoveries or theoretical innovations shall become part of the body of established science. Rather it is his professional community.
>
> (Kuhn 1963: 394–395)

Although Kuhn here speaks of the individual's 'discoveries' becoming, or not becoming, part of the body of established science, strictly (and according to his own analysis) it is only the assimilation into science that makes it proper to speak of them as discoveries in the first place. Those that are rejected, are not rejected discoveries, they are non-discoveries. And those that are accepted are not simply discoveries that are being confirmed, they are being

constituted as discoveries in the course of being accepted. Kuhn is here using the manner of speaking which collapses the different phases of the process of discovery into one moment. The burden of his argument, of course, goes in the opposite direction and is explicitly designed to open up for us the historical structure of scientific discovery.[6]

Historical work on scientific innovation has not only deepened our awareness of its complexity, it also explains a certain blindness to its social nature. One of the long-standing themes circulating in our culture is the anti-social character of genius. The genius, like the prophet, is typically represented as withdrawing from society and speaking to it from the outside. Creativity, like the prophet's authenticity, is associated with isolation and asceticism. Stories and myths celebrating the themes of detachment, otherworldliness, and sartorial and culinary indifference surround persons of genius. They can be traced from the ancient philosophers, through the heroic figures of the scientific revolution such as Boyle and Newton, to the present day (cf. Shapin 1991). Wittgenstein himself is still the focus of such stories. So we need to be on guard. These cultural stereotypes will give a wholly undeserved credibility to the idea that creativity is, or can be, or must be, an individual accomplishment.

The community decides whether the deviant behaviour of the would-be rule follower is error, or confusion, the misapplication of existing rules, or the innovative following of a new rule. The innovator doesn't earn the description 'following a new rule' until the community decides to award the title – at which stage it becomes a shared institution. While the issue is under negotiation competing definitions will be voiced. The innovator will presumably wish to call his own behaviour, 'following the new rule R*', rather than 'misapplying the old rule R', and, in so calling it, will be making a small contribution to constituting it as this. It would be naive to suppose this definition of the situation is in any way privileged. Even if the candidate innovation becomes accepted as an innovation, it is not a forgone conclusion that it will be defined in the terms proposed by the innovator. The community might celebrate the innovation as the discovery of R**, not R*, even if R** was never brought into consideration by the innovator, or would have been rejected if it had. There are many examples from the history of science to remind us of this possibility. Joseph Priestley is sometimes celebrated as the discoverer of oxygen, a substance he didn't believe in. He thought he had discovered dephlogisticated air.

Gregor Mendel is often spoken of as the discoverer of Mendel's laws. A closer look at the record, however, has suggested to historians of science that Mendel was no Mendelian, i.e. not a Mendelian in our sense. Again, Robert Boyle is famous for Boyle's law, but nothing like the modern formulation is to be found in his writings.[7]

Kuhn's sociologically oriented history, and work recognisably in this tradition, might be relied upon to provide support for a collectivist account of rule following. This very expectation can lower its value as evidence. To offset this effect it will be interesting to ask where an authority such as Karl Popper stood on this question? On certain methodological issues Popper was a self-professed individualist, so we might expect his account of innovation in science to lend support to McGinn. The opposite is the case. Popper discussed Robinson Crusoe in his *The Open Society and its Enemies* (1969). He began with a thought experiment. Suppose a clairvoyant produced a book by dreaming, or automatic writing. This book turns out to be identical to one produced years later by a great scientist, who had never seen the clairvoyant's book. Popper asked: 'is it advisable to say that the clairvoyant produced a scientific book?' (P: 219). He gave a negative answer: it is not advisable to treat the clairvoyant's book as scientific. A scientific work is one resulting from the scientific method. The clairvoyant's production does not result from the application of scientific method, so it isn't a scientific work. Popper said we might call it 'revealed science'. He then applied these considerations to Robinson Crusoe. He assumed that Crusoe succeeded in building laboratories and observatories on his island, and that he wrote a great number of papers based on his observations. He even succeeded in producing systems which coincide with the results of present-day scientists. Now, asked Popper, is this real science – or is it revealed science? Acknowledging that some would say the former, and accept Crusoe's science as real, Popper nevertheless asserted the latter: 'I assert that this Crusonian science is still of the "revealed" kind' (P: 219). There is, he explained, an element of scientific method missing from Crusoe's activities. There is nobody to check his results, or to criticise or argue with him. There is no friendly–hostile co-operation (P: 217–218). The lack of this '*social aspect of scientific method*' (P: 217) diminishes the status of Crusoe's attempts, whatever their outcome happens to be. The aspect that is eroded is the *objectivity* of knowledge. Scientific objectivity, said Popper, 'is not a product of

the individual scientist's impartiality, but a product of the social or public character of scientific method' (P: 220).

As we shall see, Popper later retreated from this admirably clear stance and formulated his views on objectivity in metaphysical language, but there is no evidence he changed his mind on Robinson Crusoe.

McGinn's intuitions lead him in the opposite direction to Popper. Not only can a Crusoe figure follow rules and innovate, the same applies to hypothetical agents placed at a far greater distance from society. McGinn brings in Romulus and Remus (M: 196) and then takes his individualism even further. Though admitting that he cannot prove it, he suggests God could have created a solitary rule follower alone in the universe for all time (M: 198). But, the collectivist will want to ask, while God could certainly create something we might unreflectively want to call a 'rule follower', is it strictly proper to describe it in this way? Should we not follow Popper and say the creation lacked the history, background or context necessary to justify the title of rule follower? The collectivist could even turn biological and functional considerations against the theological intuitions of the individualist. Millikan engaged in a theological fantasy similar to McGinn's, but came, I think rightly, to an opposite conclusion. Suppose, she says, God created a replica of ourselves, the same down to the most minute molecular details. Such a creature would have a heart and a brain exactly like ours. That, at least, is how we might at first describe the situation, but a moment's reflection is enough to raise a question. Would the replica organ really be, say, a heart? Perhaps not even God could produce a heart as an isolated entity or as an arbitrary act of creation – not, at least, a real heart, a real member of the biological category. Even God would have to work through the usual historical, evolutionary and biological channels to produce a heart, properly so called, because these are preconditions built into the definition. If God can't make a biological heart, outwith its real context, it isn't obvious that he could make a real rule follower either. Perhaps all he could do here was to make a fake or a replica, something that was as little a real rule follower as Popper's revealed science was real science.

It is unfortunate that so much philosophical discussion is conducted on the level of extravagant science fiction, myth or impressionistic theology. Does it matter what we choose to say about such hypothetical and contrived cases? Perhaps not, though it

can matter a great deal why we are inclined to say what we do. Something lies behind our responses. Popper advised us to withhold the title of real science from the Crusoe case because in doing so he signalised his sense of the importance of the public and critical character of real life science. He was right to avoid dogmatism (cf. RFM II: 62). It isn't the logic of concepts or the meanings of words that compels us in these cases – finitism assures us of that – it is more a sense of analogy based on our general strategy and preferences. I have already indicated that Wittgenstein's private language argument, though wholly reasonable and plausible as a standpoint, lacks any ultimate, compelling logical force. Its significance is best seen as methodological and illustrative. We can look upon his figure of the private diarist in PI: 258 (who undertakes to enter 'S' in his diary every day he experiences a certain, mysterious, inner and purely private 'sensation') as akin to Popper's clairvoyant. The diarist functions in the argument as an hypothetical test-case designed to elicit responses whose real significance lies elsewhere. We must always trace back what is said about hypothetical cases to the real, paradigmatic case, that is, to science as it really lies before us, or to rule following as we find it around us. These are the true objects of our concern and curiosity.

Much of the argument I have been examining in this chapter can be understood in terms of two rival syllogisms: the individualist syllogism and the collectivist syllogism. According to the individualist syllogism:

Isolates cannot participate in institutions
Isolates can follow rules
---
therefore Rules are not institutions

According to the collectivist syllogism,

Isolates cannot participate in institutions
Rules are institutions
---
therefore Isolates cannot follow rules

Let us look at each of them in turn, taking the collectivist syllogism first. If the major premise, asserting that 'isolates' cannot partici-

pate in institutions or conventions, is read as saying *physically* isolated persons cannot ever do these things, then it is false. And the same applies to the conclusion. Of course, this is not how a collectivist intends them to be read. What is meant by the premise and the conclusion is that a *socially* isolated person is disqualified. It is the individualist who typically thinks 'isolated' means 'physically isolated'. Turning to the individualist syllogism, we can see that both premises refer to 'isolates', but while one of the premises needs the qualification 'social' if it is to be tenable, the other requires the qualification 'physical'. Thus the major premise asserts that isolates cannot participate in institutions, but Crusoe's second thoughts (about the gold and silver coins) remind us that this is tenable only if 'isolates' refers to social rather than physical isolates. The minor premise asserts that isolates can follow rules, which is perfectly true, and conceded on all sides, provided it is saying physical isolates can follow rules. Individualists have not made good the claim that social isolates can follow rules because they have no account, other than a dogmatic one, of how objective standards of right and wrong can be provided in such cases. The two premises of the syllogism are therefore about different objects, one addressing the class of social isolates, the other the class of physical isolates. The individualist syllogism therefore collapses into the following form:

Social isolates cannot participate in institutions

Physical isolates can follow rules

---

therefore Rules are not institutions

which is fallacious. Both premises are true but the conclusion does not follow.

9

RULES AND THE STATE OF NATURE

The debate surrounding the social theory of rule following has a long history. The Kripke–McGinn confrontation, if I may so call it, is simply a re-run of a debate that took place in the 1780s between David Hume in Edinburgh and Thomas Reid in Glasgow. As today, the point at issue was the nature of the compulsion of rules, though Hume and Reid drew their examples from law and morals rather than arithmetic. The obligation to keep promises, rather than the necessity of addition was at the centre of the dispute. G. E. M. Anscombe has already drawn attention to the similarity between Wittgenstein's social theory of rules and Hume's penetrating and original approach to moral philosophy (Anscombe 1969; 1978b). I shall take up the same theme, but expand on Anscombe's suggestion by a more detailed study of the relevant doctrines of Book III of the *Treatise*. I shall also show how the comparison can be deepened by bringing in Reid, in his role as Hume's opponent. In this way, the positions of both sides in today's debate can be mapped onto those of their eighteenth-century Scottish forebears.

Hume's declared intention on the title pages of his *Treatise* was to introduce the experimental method into moral enquiries. One way he did this was by thought experiments. He studied the obligation to keep promises by imagining someone who breaches all our expectations by feigning ignorance of this moral requirement. This person borrows money, and then 'naively' asks why he should return it. Just as Kripke's sceptic keeps asking for the fact that constitutes meaning addition, and the awkward pupil of *Investigations 185* wants to know why he must write 1002 after 1000, so Hume's trouble-maker wants to know what this fact of 'having to' pay back his loan consists in. As Hume put it:

> I suppose a person to have lent me a sum of money, on condition that it be restor'd in a few days; and also suppose, that after the expiration of the term agreed on, he demands the sum: I ask, *What reason or motive have I to restore the money*? (T: 479).

Like Kripke and Wittgenstein, Hume posed a question that was both simple and precise. The circumstances of following a rule such as 'add 2', and of returning a sum of borrowed money at an agreed time are quite clear. The issue isn't whether this or that is required, it is about the character of requiredness. Hume and Kripke and Wittgenstein use examples which, once our taken-for-granted expectations are breached, take on a puzzling, almost mysterious air. How can we produce some sign or formula – the promise or the rule – and thereby *bind* ourselves in certain ways? What is this necessity that then grips us?

Because the sceptical challenges discussed by Hume, Kripke and Wittgenstein call into question what everyone takes for granted, they always have an 'obvious' answer. The obvious answer to Wittgenstein's pupil is that the rule requires us to write 2004, and the obvious answer to Kripke's sceptic is that we just *know* what we mean. The obvious answer to Hume's question is that we should return the money out of a 'regard to justice'. This reply will be sufficient 'if I have the least grain of honesty, or sense of duty and obligation'. Hume, like Kripke and Wittgenstein, didn't let the matter rest there. He applied critical pressure by imagining someone to whom the appeal to duty and honesty meant nothing. He supposed we were confronted by someone in the 'state of nature', i.e. someone in a 'rude and more *natural* condition', who has not been trained up in a particular moral code: 'For one in that situation wou'd immediately ask you, *Wherein consists this honesty and justice, which you find in restoring a loan, and abstaining from the property of others?*' (T: 480).

The phrase 'wherein consists' was Hume's way of posing the question: 'what is the fact of the matter', the very question Kripke's sceptic asked about meaning. So the question is: what is the fact of the matter that constitutes duty? This question did not arise from any assumption on Hume's part that men are naturally selfish. He didn't think they are, at least, not totally. Charity and generosity are among our natural virtues. The point is that if someone returned money they had borrowed and felt they were acting out of, say,

generosity, we would think their attitude misplaced. We would say: it isn't a question of kindness, but of duties and rights. Hume would have agreed. He emphasised the difference between acting on the basis of feeling, and acting on the basis of duty. Feelings obey different laws from duties. Feelings wax and wane, decay with time and distance, are focused on friends and family, and divide us from those outside our immediate circle. Duties have a more uniform, impersonal character to them. Our natural generosity, said Hume, is a 'confin'd generosity'. By contrast, duties, such as the duty to keep promises, is subject to no such limitation, because it is not 'the immediate offspring of any natural motive or inclination' (T: 532).

To prove this negative point about duties Hume conducted a search for the alleged 'natural motive' to keep promises. Is there perhaps an act of mind which underlies promising and makes it what it is, namely an utterance that places us under an obligation? Let us, said Hume, 'run over all the faculties of the soul, and see which of them is exerted in our promises' (T: 516). The candidates he considered and dismissed were: resolving, desiring, and willing. Promising isn't the same as resolving, because a mere resolve carries no obligation. A resolve we make today, we are free to modify tomorrow. Nor is promising the same as desiring, because we often have no desire to do what we have promised, but we are still bound. Is promising, then, a species of willing? This is more difficult, because a variety of different formulations need to be considered. To begin his attack Hume asserted that the will can only influence present actions. An act of will can't be a promise, because promises can reach forward into the future, and an act of will can't (T: 516). But can't we will now that we now be under an obligation? If we can now place ourselves under an obligation, and the obligation concerns a future action, we have overcome the problem: it is the obligation that reaches forward. This is acceptable provided that certain preconditions are satisfied. The obligation must now be available for us, it must already exist, if we are to be able to make use of it. And if we suppose that, we have helped ourselves to precisely the thing we were trying to explain, namely, what obligations are, and how they can be created.

This brings us to the nerve of the problem. Can an individual act of will, or a resolve, *create* the very obligation that is the object of the act of willing or resolving? Can we impose an obligation on ourselves, and thereby create the obligation we are putting ourselves under? Hume believed this made no sense: it collapses together a

mental act and its object. The mental act is the act of putting ourselves under an obligation, the object of the act is, or involves, the obligation itself towards which we are oriented. This object cannot just be the creature of the individual act if the result of the act is to serve as a genuine constraint on the individual. The object must be independent of the act, but the idea that we can create an obligation by a mere act of individual will fails to satisfy this condition of independence. As Hume put it, 'The will has here no object to which it cou'd tend; but must return upon itself *in infinitum*' (T: 518). In simple terms, we can't explain the creation of obligations by simply citing the resolve to create them. Such acts of will and resolution presuppose obligations. Saying, 'I am hereby under an obligation', does not, of itself, ensure that one is under an obligation. Even inventing and employing a sign which, by its definition, is said to mean 'I am hereby under an obligation' doesn't settle the question of whether, after its use, you really are under any obligation (cf. Anscombe 1969). And it certainly doesn't settle the further question of the nature and origin of such obligations.

Hume formulated this negative result in a striking and imaginative fashion. It showed, he said, that: 'a man, unacquainted with society, could never enter into any engagements with another, even tho' they could perceive each other's thoughts by intuition' (T: 516).

No matter what degree of resolution underlies an individual's words, and no matter how transparent their sincerity to the other party, a 'man unacquainted with society' cannot bind or commit himself. Moral obligation can't exist in a state of nature where each person can only rely on and mobilise their own, individual resources. So where do obligations come from, if the natural powers of the individual mind cannot generate them? Hume's answer was that they came from social conventions. They are the cumulative product of primitive and embryonic bargains struck up between individuals. Given what he has just said about promises, Hume doesn't and can't take such bargains to be contractual relations or things secured by promises. His task was to build up the conditions where people *can* bind themselves by their words, starting from circumstances where they can't. What resources are available to make this construction? In the first instance the answer is: our natural, psychological tendencies and faculties. An obvious candidate would be our natural generosity and willingness to help others, and co-operate with them. But in as far as our relations with one another are explicable by these traits, we don't need promises.

Putting one another, and ourselves, under moral obligations is something that comes into play at the very point when these amiable traits can't be relied upon. In any case, in Hume's view, such natural traits are of limited scope. It is not, he said, our natural virtues alone that explain conventions; it is our interests and our inductive capacity to frame reasonably well-informed expectations about the behaviour and reactions of others.

The problem is encapsulated in a famous example. Suppose, said Hume, that your corn is ripe today and mine will be ripe tomorrow. It is profitable to both of us that I should help you today, and you should help me tomorrow. You have no particular feeling of kindness to me, nor I for you; so if I help you, how do I know you will return the favour? What, then, can move us into co-operation and prevent both of us losing our harvests 'for want of mutual confidence and security' (T: 521)? Hume's solution was that a bargain can be struck only when I can contemplate doing my part in the knowledge that it is strongly in your interests to return the service, and in the knowledge that you know it is in your interests to reciprocate, and are likely to act on that knowledge. Under these circumstances you can enter into an engagement to do your part, because you can demonstrate to me that you have something to lose by defaulting, and something that will outweigh the gain. In this way self-interest, which seemed to work against the possibility of a bargain, can be turned into a device for showing there is no live option but to fulfil the bargain (T: 492). When one such bargain has succeeded, that fact enters into the calculation behind the next exchange. Trust can begin to emerge on the basis of successful exchanges not themselves based on trust. A history of such small, low-risk acts of successful co-operation can then provide the basis for larger, longer-term, co-operative enterprises. Once a string of such acts of co-operation has been established we have what Hume called a 'convention'. A convention, on this understanding of the term, is *not* a habit or a custom, though it might become habitual or customary. Hume carefully distinguished habits and customs from conventions, on the grounds that a notion like 'habit' can't explain the historical emergence of a sequence of actions. Something else, such as the calculations that grow out of self-interest, must first of all create and sustain the sequence in order to make it available to become a habit.

We can now see why the 'man unacquainted with society' was powerless to make a bargain. Whatever his inward tendencies, there

was nothing in his surroundings he could use to impose a cost on himself if he failed to perform his part of the bargain. His problem was not that he had nothing to offer, but that he had nothing to lose. He could not demonstrate to the other the cost of non-compliance in such a way that compliance would be confidently predicted. Being powerless to lose his freedom of action, he could not bring himself under an obligation. As Hume made clear, the mechanism for incurring costs is social. They involve the predicted behaviour of other people in possible future bargains. To create obligations we must exploit our known involvement in, and known dependence on, other bargains – i.e. our need for a 'correspondence of good offices' (T: 521).

Promises depend, in the way just described, on conventions, but it would be wrong to express Hume's theory by saying that we keep promises because there is a convention to keep them. We keep them, so far as we do, because we are under a moral obligation to keep them. Hume wasn't denying this. He was saying the obligation is created by, is grounded in, and is explicable by, the convention. The convention itself just is the pattern of co-ordinated, calculative behaviour, the pattern of reciprocal help, or the correspondence of good offices. The end to which the convention is directed is not promise keeping, any more than our promises are promises to keep promises. Only when various conventions are emerging, or have emerged, is it possible to send signals able to perform the role of promises. The words 'I promise' only carry information when they are uttered in circumstances where both the sender and the recipient of the signal can think that a train of events is being set in motion that is outwith their control, and is of the kind that will secure compliance. Only then will the words be more than mere words, and constitute a genuine move in a bargaining game. This is why Hume declared: 'interest is the *first* obligation to the performance of promises' (T: 523).

Can this theory explain the specifically moral character of the obligation to keep promises, over and above its merely prudential aspect? Hume was ready for this question, and drew on the resources of the theory of moral sentiments to provide an answer. A new convention is a new object, and a new object provides a new stimulus for our sentiments. In general, conventions prompt feelings of approbation – otherwise we would not exert ourselves in ways that sustain them. So when interest has done its work 'a sentiment of morals concurs with interest, and becomes a new obligation

upon mankind' (T: 523). Hume was saying a convention generates a phenomenological change. We feel our new, socially structured predicament in a characteristic way. The power others come to have over us, ultimately because of the interlocking of our own and their interests, has a specific quality of feeling. We call it 'moral obligation'. The emergence of this sentiment of morality is, for Hume, an entirely natural process: 'After that interest is once establish'd and acknowledg'd, the sense of morality in the observance of these rules follows *naturally*, and of itself . . . ' (T: 533). This new lustre of sentiment attaching to our conventions can obscure the underlying interests and calculations from our view. We see people making resolutions; then we see them under obligations, so we believe ourselves to have seen the resolution creating the obligation. Thus:

> we cannot readily conceive how the making use of a certain form of words shou'd be able to cause any material difference. Here, therefore, we *feign* a new act of the mind, which we call the *willing* an obligation; and on this we suppose the morality to depend (T: 523).

It was this illusion that made it necessary for Hume to begin his argument by attacking the theory that obligations are caused by promises, rather than being presupposed by them.

We keep promises because it is our duty to do so, and the same applies more generally to our respect for other people's property. We are loyal to governments and monarchs because it is our duty to obey. But for Hume the process of doing our duty because it is our duty, is problematic. He couched the problem in terms of acting 'out of regard for the virtue of an action'. An act, he said, is virtuous if it has a virtuous motive. We are not content, morally, with superficial, merely outward, compliance. We require a special inner orientation for an act to be genuinely virtuous. This was Hume's version of the requirement of conscientiousness. Can that inner motive be a sense of duty? The trouble with that answer, said Hume, is that to act out of regard for the virtue of an act presupposes the act *is* virtuous. How then did it become virtuous? If its virtue depends on its motive, and the motive depends on its virtue, we have gone in a circle. There must, therefore, be something else capable of setting the system in motion. This is why he insisted that interests must work alongside regard for virtue. Once a conventional morality is established then it becomes possible to act purely out of a sense of duty, though such purely conventional forms of obser-

vance may strike us as somewhat hollow. It offends our sense of natural virtue when people are merely moral conformists.

In Hume's words the argument was that:

> all virtuous actions derive their merit only from virtuous motives, and are consider'd merely as signs of those motives. From this principle I conclude, that the first virtuous motive . . . can never be a regard to the virtue of that action, but must be some other natural motive or principle. To suppose, that the mere regard to the virtue of the action, may be the first motive, which produc'd the action, and render'd it virtuous, is to reason in a circle. Before we can have such a regard, the action must be really virtuous; and this virtue must be deriv'd from some virtuous motive: And consequently the virtuous motive must be different from the regard to the virtue of the action (T: 478).

The original motive is self-interest and our natural feelings (T: 478 and 479). The overlay of moral approbation that attaches itself to the resulting conventions is also reinforced by sympathy if we can see a public good furthered by it (T: 499–500).

Thomas Reid provided the perfect foil for Hume. Conscientious, respectable, forthright but as subtle as he was tenacious, Reid worked quite deliberately from within the very normative system Hume's investigations rendered problematic. Reid was, however, no naive victim of Hume's sceptical challenge. He was a representative of a definite, identifiable tradition of self-reflection on our everyday practice, namely the Scottish philosophy of common sense (see Davie 1991, 1994; and MacIntyre 1988). This tradition, not the practices themselves, was the true object of Hume's scepticism. Reid's discussion of 'the nature and obligation of a contract' in his *Essays on the Active Powers of the Human Mind* (1788) began with the words:

> The obligation of contracts and promises is a matter so sacred, and of such consequence to human society, that speculations which have a tendency to weaken that obligation, and to perplex men's notions on a subject so plain and so important, ought to meet with the disapprobation of all honest men (AP: 435).

He went on to identify Hume's speculations as an instance of this deplorable kind.

Reid distinguished what he called the 'solitary' from the 'social operations of the mind'. Solitary operations include seeing, hearing, deliberating, and forming and executing purposes. These can all be done 'without intercourse with any other intelligent being' (AP: 437). This distinguishes them from other operations which 'necessarily imply . . . some other intelligent being who bears a part in them' (AP: 437). These are the social operations. Reid instanced: asking questions, testifying a fact, giving a command to a servant, promising and entering into a contract. A 'remarkable distinction' between the solitary and social operations is that in the social operations verbal expression is essential. According to Reid, philosophers commonly assume the social operations may be resolved into the solitary. When they list the mental faculties they therefore tend to omit the social ones. Reid rejected this form of reductionism: 'The social operations', he said, 'appear to be as simple in their nature as the solitary' (AP: 439). The capacity to perform the social operations belongs to each individual prior to the growth of reason; hence they cannot be the result of calculation. We do not need to be taught to ask or refuse or threaten or supplicate. Our capacity to do these things belongs to the human constitution: they are evinced and understood by a kind of 'natural language' of gesture and expression. This natural language is the foundation for fully-fledged verbal intercourse (AP: 440). Even brutes have some ability to communicate in this way. But there are two social operations, according to Reid, no brute animals can perform. 'They can neither plight their veracity by testimony, nor their fidelity by any engagement or promise' (AP: 441–2). This is a human prerogative; it is fundamental to human nature, and the basis on which society rests. 'Without fidelity and trust', said Reid, 'there can be no human society' (AP: 443).

The stage was now set for the confrontation with Hume. Whereas Hume wanted to ground the obligation to keep promises in established patterns of social behaviour, Reid insisted these social processes themselves presuppose the moral capacities and commitments Hume wanted to explain. The conflict was epitomised by Hume's assertion, and Reid's denial, that it was necessary to look beyond the sense of duty in order to explain morality. For Hume it was obvious we can do X out of a feeling that X ought to be done, if, and only if, we *already* accept that X is a virtue. Explaining our doing X by reference to duty can therefore do nothing to illuminate why X, rather than *not X*, is a virtue. Hume's aim was to address

this question while avoiding pseudo-explanations. To avoid vacuous accounts of promise-keeping, or any other virtue, it was necessary to find something to support the entire moral system of promise-giving, promise-keeping obligations, *and* our sense of duty to keep promises. To locate such a thing it was necessary to step outside the system – hence the appeal to an underlying structure of interests.

Reid saw no circle and no vacuity here. For him there was no need to go behind our sense of duty, because the faculty of conscience by which we perceive it, is grounded in our constitution. It was the attempt to do what Hume did, and to locate another motive, that was logically questionable. To appeal to interests meant, for Reid, introducing a factor running completely counter to our moral motives. Interest and morality were thought of, by Reid, as being wholly antithetical (except, of course, for the special case of the interest in morality itself). Reid developed this point, not about promises in particular, but about Hume's general claim that no action can be morally good 'unless there be in human nature some motive to produce it, distinct from the sense of its morality' (T: 479). To Reid this was 'shocking'. Consider a judge who passes sentence without regard for anything other than the justice of the case. Isn't this obviously right and good? But didn't Hume's principle make this judge immoral, because he has no motive distinct from morality? Surely everybody could see Hume had got things back to front. Virtuous actions, said Reid,

> are so far from needing other motives, besides their being virtuous, to give them merit, that their merit is then greatest and most conspicuous, when every motive that can be put in the opposite scale is outweighed by the sole consideration of their being our duty (AP: 448).

Reid acknowledged what he called 'collateral inducements' to virtue, e.g. that it requires no artifice and dreads no detection. Reflection tells us virtue carries these rewards, but according to Reid, we do not arrive at our belief in the virtue of justice, or the virtue of keeping promises, by any such calculation of advantages. The notion of right and wrong, said Reid: 'is discerned, not by any train of reasoning, but by an immediate perception (AP: 444).

When Hume carried out his survey of the various faculties of the soul to see which was implicated in promising he must, therefore, have passed over the true candidate. Hume correctly eliminated resolving, desiring and willing, but failed to see that the faculty of

Reason was responsible. Reason, for Reid, was the faculty by which we grasped moral truths and formulated moral judgements:

> Every determination of the understanding, with regard to what is true or false, is judgement. That I ought not to steal, or to kill, or to bear false witness, are propositions, of the truth of which I am as well convinced as of any proposition in Euclid. I am conscious that I judge them to be true propositions; and my consciousness makes all other arguments unnecessary (AP: 464).

Hume was simply looking in the wrong place: he had missed the obvious fact that our reason gives us insight into the facts of moral obligation.

To drive home the Euclid-like certainty of the obligation to keep promises, Reid tried to justify the principle on grounds of logical coherence alone. Recall that in Hume's example, a representative of the state of nature claimed to be unaware of any obligation to return property he had borrowed. Hume asked us to suppose the borrower found the appeal to his sense of duty 'perfectly unintelligible and sophistical'. In Reid's eyes this was incoherent. First, Reid observed that either the 'rude man' is right, or he is wrong. If he is right, he must have perceived a weakness infecting all promises. How can a promise have no force in the state of nature, but carry genuine obligation for people trained according to a certain discipline and education? Hume's example either shows the triviality that savages can make mistakes, or it destroys all promises. Reid's second attack began with the observation that Hume gave no real examples of the kind of episode he described. For Reid this was no surprise, because all societies are founded on man's fidelity. Behind this factual observation Reid perceived a deeper point: it doesn't make sense to suppose a society where people borrowed but could not understand the obligation to pay back. The very idea involves a contradiction:

> That a man should lend without any conception of his having a right to be repaid; or that a man should borrow on the condition of paying in a few days, and yet have no conception of his obligation, seems to me to involve a contradiction (AP: 451).

'Lending' surely means 'transferring property whilst retaining ownership, and hence the right to reclaim'. If we describe someone

as 'lending' and then say they have no conception of their right to be repaid, we are open to the charge of having misdescribed the event. Shouldn't we have said they were simply giving? This suggests there is some kind of logical connection between lending and the right of repayment, and between borrowing and the obligation to repay. The connection is what philosophers after Reid called an internal relation, i.e. a relation in which the terms cannot be adequately specified independently of one another.[1]

Reid set these arguments in a theological framework, and it is important to make clear where he stood on the question of God's relation to the fundamental truths of morality. The reason for being sensitive to Reid's theological stance goes beyond that of historical accuracy and the desirability of understanding a thinker in his own terms. It will prove important when we come to position the Hume–Reid controversy in relation to today's discussion of rule following. Reid addressed the theological dimension in the very last passages of *Active Powers*, immediately after the arguments described above. Our nature, he said, is a faint but true copy of God's nature. We have a faculty for conceiving moral truths; can we therefore doubt that God also has such a faculty, albeit more perfect than ours? There are moral truths, and: 'these truths cannot be hid from him whose understanding is infinite, whose judgment is always according to truth, and who must esteem every thing according to its real value' (AP: 481).

These are not just pious formulae. They contain a very precise theological and metaphysical message for those – such as Reid and his contemporaries – whose ear is correctly tuned. His audience would be scanning Reid's text for indications as to which of God's attributes are the ones deemed most worthy of respect and reverence. Was it, for example, God's will, or was it his foresight? In the event we find Reid saying God is connected to goodness, not through the arbitrary decrees of his all-powerful will, but through his divine understanding of moral truths. God has a faculty for detecting these truths, but it is, of course, a perfect faculty – so none of them can remain hidden from him. God wills the good, and so he wills what he sees to be good. In saying this Reid made a consequential choice, and sent a message of great significance to his readers – a message further underlining his divergence from Hume.

Similarities between the Hume–Reid debate and today's discussion of rule following will have become apparent, but for

completeness I shall provide a summary of the main points. A preliminary difficulty, however, needs to be removed. Clearly the obligation to keep promises is a rule regulating the interaction between two or more persons. From the outset the rule has a social reference, but nothing in the comparison will hinge on the fact that the rules discussed by Hume and Reid had a social object. We must distinguish the content of the rule from the nature of the rule, and the object of the necessity it imposes from the character of that necessity. The comparison will only concern the sources of compulsion rules possess, not their particular subject matter. Now for the similarities.

The first and most obvious point of comparison is that both the old and the new debates involve a sceptical challenge. Unfortunately many commentators have responded to that characterisation with quite excessive sensitivity, denying that Wittgenstein's work should be seen as sceptical. In order to avoid terminological disputes about what is, or is not, an exercise in 'scepticism' it may be useful to characterise the common factor in another way. I shall say both debates utilise 'breaching experiments' – a technique well-known to sociologists. A breaching experiment is a deliberate flouting of a norm in order to make its operation and significance visible and open to study. For example, if you stand very close to a stranger you rapidly discover that there are norms that define being 'too close'. You could do this deliberately to find out the variables governing the norm, e.g. age, sex, and nationality. Pretending to be naive, and hence questioning what people take for granted, is another version of this technique and is used as a literary device by Hume, Kripke and Wittgenstein. All three writers use hypothetical breaching experiments. A typical response to a normative breach is exasperation or anger. Reid's response to Hume had something of this character, though of a controlled kind. 'It is not the man I impeach, but his writings. Let us think of the first as charitably as we can . . . ' (AP: 435). Wittgenstein perceptively built these feelings of exasperation into his literary presentation.

> Now we get the pupil to continue a series (say +2) beyond 1000 – and he writes 1000, 1004, 1008, 1012. We say to him: 'Look what you've done!' – He doesn't understand. We say: 'You were meant to add *two*; look how you began the series!' – He answers: 'Yes, isn't it right? I thought that was how I was *meant* to do it' (PI: 185).

A second comparison is that both Hume and Kripke resorted to a systematic, eliminative search through the contents of the individual mind to find the fact constituting the compulsion of a rule. Both presented the search as ending in failure. Wittgenstein did not engage in a set-piece search of this kind though it is an acceptable and ingenious way to summarise his argument. Wittgenstein did, however, share with Hume and Kripke the dramatic device of presenting the conclusion to their argument in terms of what we could, or could not, see if we could look into the mind. They all invoked the idea of the transparency of the mind. Wittgenstein and Kripke imagined what God might see. Hume supposed that two persons encountering one another in 'the state of nature' were transparent to one another. The point was that such transparency would still not reveal the locus or source of obligation. Individual thoughts are no different from individual utterances, they are just one sign replacing another sign (Z: 36 cf. Z: 557–560).

The third comparison is that in both debates the critics of the (explicit or implicit) process of elimination adopt the same response. They claim the search did not exhaust the content of the individual mind. We are looking in the right place, but for the wrong kind of thing. Like Reid before them, McGinn and other individualists say the individual mind does contain the relevant fact, i.e. the fact of necessity or obligation, but it is an irreducible fact. Its unique nature differentiates it from any of the previous candidate facts that have been considered and rejected. The only difference between today's version of this reply, and Reid's, is that contemporary philosophers are sometimes less candid, or less clear, about the denial of naturalism implicit in their views. And they certainly cannot match Reid's rhetorical style. As Reid put it:

> if the notion of duty be a simple conception, of its own kind, and of a different nature from the conceptions of utility and agreeableness, of interest or reputation; if this moral faculty be the prerogative of man, and no vestige of it be found in brute animals . . . if to be governed by it be the glory of man and the image of God in his soul . . . I say, if these things be so, to seek the foundations of morality in the affections which we have in common with the brutes, is to seek the living among the dead (AP: 404).

In both the eighteenth- and twentieth-century versions of the debate, those who perform the breaching experiment do so in order

to reveal the social, rather than the psychological or individual, character of the phenomenon under study. In both cases the opposition to the sociological approach takes the form of an irreducibility thesis and a corresponding commitment to a form of meaning determinism. In the earlier argument promising just is something that we can do, and the obligation derives from intuiting the meaning of the commitment. In the later argument, we simply know what we mean by a concept or rule, and the corresponding necessity derives from its meaning. The causes and effects which the collectivist theory distributes amongst a set of social interactions, the individualist account concentrates in a mysterious act of mind with quasi-magical or divine properties.

A number of further points could be deployed to reinforce these comparisons. For example, Hume's account of the phenomenology of obligation, of the way we 'feign' an act of mind to explain to ourselves the social processes we do not immediately comprehend, is clearly an ancestor of Wittgenstein's account of how rule following (misleadingly) 'strikes us'. We are tempted into myth making because we do not know our way around the workings of our own language games and social practices. Hume's phenomenology of feigning is just Wittgenstein's bewitchment of our intelligence by language (PI: 109). If we were to widen our horizons, yet further similarities between Hume and Wittgenstein would emerge. For instance, Hume advocated a form of finitism, roundly declaring that, 'The capacity of the mind is not infinite' (T: 39). The theme emerged in Book I of the *Treatise* in his discussion of infinitary ideas in geometry, in particular in the account of parallel and intersecting lines (T: 52). Its counterpart in the moral philosophy lies in Hume's preoccupation with demonstrating the complexity of legal and moral principles. A sentiment of obligation attaches to them, but the practical application of these principles cannot flow smoothly and automatically from their meaning. They are opaque to reason and do not answer to any innate or universal tendency in our minds. This opacity derives from the local contingencies of their use, particularly the interests impinging on them. Thus:

> 'tis absurd to imagine, that in every particular instance, these sentiments are produc'd by an *original* quality and *primary* constitution. For as the number of our duties is, in a manner, infinite, 'tis impossible that our original instincts should extend to each of them, and from our very first

> infancy impress on the human mind all that multitude of precepts, which are contain'd in the compleatest system of ethics (T: 473).

This is the counterpart of Wittgenstein saying, when you give the order 'add 2', you cannot conjure up in your mind each of the steps you intend should be taken in obedience to the rule. Reid, by contrast, refers without qualms to man's 'unlimited power of conception'.[2] Today's individualists exhibit the same tendency as Reid, with their willingness to invoke infinitary capacities with which we can grasp concepts like addition in all their scope (cf. M: 162).

Individualists read Wittgenstein as if he were a latter day Reid. I shall draw my final reason for rejecting this identification from what may seem, at first sight, a strange source: Reid's theology. I shall exploit the fact that Wittgenstein, though an atheist, was inclined to express his philosophical ideas in theological terms. This allows us to locate a deeply revealing difference of opinion between them that they both expressed in the same idiom. Norman Malcolm drew attention to this interesting, and quite self-conscious tendency of Wittgenstein's to use religious language (cf. Malcolm 1993). One instance of this, though not mentioned by Malcolm, occurs in connection with Wittgenstein's reaction to Moritz Schlick's *Fragen der Ethik* (first published 1930, English translation, Schlick 1939). His response was recorded as follows:

> Schlick says that theological ethics contains two conceptions of the essence of the Good. According to the more superficial interpretation, the Good is good because God wills it; according to the deeper interpretation, God wills the Good because it is good.
>
> I think that the first conception is the deeper one: Good is what God orders. For this cuts off the path to any and every explanation 'why' it is good, while the second conception is precisely the superficial, the rationalistic one, which proceeds as if what is good could still be given some foundation.
>
> (Waismann 1965: 15)

This is a most important statement, bringing out Wittgenstein's impatience with 'rationalistic' approaches which seek to provide some foundation in reason for all natural or supernatural phenomena. On the rationalistic view, even God's commands are to be explained in terms of their conformity to a pre-existing standard,

hence God wills the Good because it is (already) good. For Wittgenstein this commitment to a rational grounding for everything was superficial. It failed to confront the ultimate groundlessness and contingency of reality; it accepts nothing as brute fact, but wants to surround everything with a halo of rationality (cf. OC: 166). A deeper, more rigorous sensibility sees this for what it is – a hopeless illusion. Reasons have got to give out.[3] As a symbol of this unavoidable, rationally opaque end-point, Wittgenstein, seeing the problem from a religious point of view, instances the commands of God: '*Gut ist, was Gott befiehlt*' – 'the good is what God commands'. 'If any proposition expresses just what I mean', he added, 'it is: Good is what God orders' (Waismann 1965: 15 and 1979: 115).

In his *Logic and the Basis of Ethics* (1949), the logician A. N. Prior called this position 'theological naturalism', to distinguish it from 'theological rationalism' (1949: 100). Reid was a theological rationalist, and theological rationalism was one of Hume's targets. Hume identified for attack the doctrine that 'the immutable measures of right and wrong impose an obligation, not only on human creatures, but also on the Deity himself' (T: 456). Wittgenstein deliberately aligned himself with theological naturalism. The position Wittgenstein adopted is also known as 'voluntarism'. One (correct) observation theological rationalists always make in disputes with voluntarists and naturalists is that if you identify 'X is good' with 'God commands X', you cannot go on to commend or praise God's act of will as itself good or righteous. To do so merely amounts to the truism that his will is his will, and God commands what he commands. Wittgenstein, however, was the last person to balk at this consequence of his position. It was a characteristic of his philosophy that he repeatedly warned against applying words that operate within a language game to the game itself. Expressions of belief and certainty apply within a taken-for-granted framework, which cannot itself be said to be 'known with certainty' (cf. PI: 679). For example, he tells us it is wrong to say, 'I know what I am thinking' (PI II: p.222). 'It is', he said, 'as if "I know" did not tolerate a metaphysical emphasis' (OC: 482). This theme crops up time and again in different guises, from his comments about the standard metre in Paris being the one thing we can't say is, or is not, one metre long (PI: 50), to his assertion that the rules of inference of a logical system cannot be called right or wrong (RFM VII: 30). If the truth is what is grounded, he said,

then the ground itself is neither true nor false (OC: 205). Seen from a religious point of view, these are all ways of accepting that you cannot commend God's evaluations or commands. Wittgenstein did not see this family of related consequences as calling into question his approach. He treated them as following from it, and hence as things to be embraced.

It is revealing to look at how we should respond to sceptics in this light. Wittgenstein told us we can dismiss sceptics, like Kripke's sceptic, but we cannot dismiss them with reasons, i.e. on the grounds that we know they are wrong. We cannot be said to know this.

> The queer thing is that even though I find it quite correct for someone to say 'Rubbish!' and so brush aside the attempt to confuse him with doubts at bedrock – nevertheless, I hold it to be incorrect if he seeks to defend himself (using, e.g., the words 'I know') (OC: 498).

This passage shows just how apt Kripke's label of 'sceptical solution' was for Wittgenstein's position, providing 'sceptical' means the rejection of any pretensions to ultimate justification (cf. also RFM VI: 29).[4]

I need to relate this point to my previous claim to have given a straight, sociological answer to Kripke's sceptic, and to have done so in Wittgensteinian terms. In giving that answer I treated the sceptic as posing reasonable questions such as: how do we mean things by our words? and, how do we follow rules? These are legitimate enquiries and deserve something better than sophisticated evasions. But how can such a stance be rendered consistent with Wittgenstein's comments about dismissing the sceptic? Obviously, it depends on what the sceptic is taken to be doing. Let us go back for a moment to Jones the property owner who faced the sceptical challenge over what constituted the fact of his ownership. I said the challenge was a legitimate request for explanation, but there are different ways of 'hearing' such a challenge. Such requests are often 'heard' as implicit criticisms of what is to be explained. They are treated as rejections, as if the sceptic had said, 'There is no such thing as property!' This might reasonably be met with a brusque 'Rubbish!' Another response to 'hearing' a radical criticism might be, 'Wrong! I KNOW there is such a thing as property!' This is what Wittgenstein advised against. Both of these responses are meant as (a) moves in the language game, and (b) responses to an implied rejection of the game. I 'heard' the sceptic as saying something quite

different, something that was to be met neither by saying 'rubbish' nor by saying 'I know'. What I heard called for an overall, explanatory view of the relevant institution. And that was Wittgenstein's own response in the *Investigations* and *Remarks* where he introduced the ideas of language games and invoked the notion of institution.

For those of us accustomed to secular habits of mind and speech, Wittgenstein's tendency to express things in a religious idiom is in danger of seeming merely quaint. We hear him say, for instance, 'a good angel will always be necessary, whatever you do' (RFM VII: 16), and it means little to us. It should, however, be resonating with the doctrines of meaning finitism which make every application of a concept a contingent and, in a sense, hazardous accomplishment. Wittgenstein even resorted to a theological idiom in trying to convey some of his most characteristic and challenging thoughts about mathematics. He asked, for instance, what harm is done by saying that God knows *all* the irrational numbers? His reply was that this hides certain problems. What problems? The answer, surely, is the problems exposed by meaning finitism. Thus, referring to the infinite sequence that can be used to calculate the value of pi, he said:

> I want to say: Even God can determine something mathematical only by mathematics. Even for him the mere rule of expansion cannot decide anything that it does not decide for us.
>
> We might put it like this: if the rule for the expansion has been given us, a *calculation* can tell us that there is a '2' at the fifth place. Could God have known this, without the calculation, purely from the rule of expansion? I want to say: No (RFM VII: 41).

Notice that the issue here is God's omniscience. The claim is that there is literally nothing to be seen, no existence to be discerned, even by so discerning an eye as God's, that exists *in advance* of the calculation to be performed. Just as God is not to be thought of as possessing moral intuitions which detect a pre-existing moral rule, so, for the finitist, He cannot be thought of as following a mathematical rule by tracing out a pre-existing mathematical reality. Voluntarism is the theological counterpart of finitism. Wittgenstein's talk about God is best understood as encoded talk about society, so God's position is just like our (collective) position in this respect. He must create meaning as he goes along, just as we

must. The rule cannot decide anything for him that it cannot decide for us (cf. RFM V: 13).

Because of the centrality of these themes for Wittgenstein's philosophy, we cannot shrug off his theological utterances as merely decorative, or fanciful. They are, of course, rhetorical devices, but highly significant and subtly structured ones, calling knowledgeably on the resources of a rich cultural tradition. As a passionate reader of Tolstoy, Kierkegaard and Augustine, we know that Wittgenstein was well informed about this tradition. Those of us who are less well informed, or whose sensibilities are blunted, must take care, lest we miss the message. Wittgenstein the theological naturalist or voluntarist, was deeply opposed to the rationalistic standpoint represented by Reid in the Hume–Reid debate. It was Reid, not Hume, who was seeking to explain everything by reference to an underlying reason. For Reid it was God's omniscience, not His all powerful will, that was important, and this standpoint informed Reid's entire position about the nature of rules and obligations. For the rationalist, God is constrained by pre-existing rules; for the voluntarist He cannot be so constrained because He creates the rule in a sequence of arbitrary acts of will. The former position is the theological expression of meaning determinism, the latter has already been identified as the analogue of meaning finitism. Hume dispensed with God in his explanations, but that fact alone would still have left him free to adopt a humanistic rationalism, or pseudo-naturalism, in which an alleged human capacity for insight into rational necessity and rational compulsion was held to be the driving force in his explanation. Of course, Hume did not reject theological rationalism merely in order to replace it with any such humanistic rationalism. Like Wittgenstein, he rejected all forms of rationalism and meaning determinism in favour of a thorough-going naturalism and conventionalism. Society, for these two thinkers, certainly took over from God but, and this is the vital point, from a voluntarist God. This, ultimately, is why we must assimilate Wittgenstein to Hume's side, not Reid's side, of the controversy.

The Humean comparison provides a good way to summarise Wittgenstein's position. We can say: for Wittgenstein rule following is an artificial virtue, i.e. a capacity grounded in convention. The mistaken comparison with Reid would reduce rule following to one of the Active Powers of the individual human mind. We can also identify the individualist's mistake as confusing the artificial with

the natural virtues. In some respects Hume's work, though not confronting such hard cases as Wittgenstein's, is superior in clarity. Hume offered a well defined model of conventional behaviour, and also realised that its strategic, and hence self-referential, character demands an account of its origins. His response was to describe a possible mechanism of emergence, i.e. a method of 'priming' the system, based on the individual ability to interact with others calculatively and in the light of interests. As history the story may be questionable, though as a continuing mechanism for sustaining conventions it is surely illuminating. It can be counted as a philosophical fiction, perhaps individualist in its temper and method, but certainly not in its consequences. Wittgenstein can be seen as Hume without the origin myth.

Could Wittgenstein's lack of a corresponding model of an institution be defended in terms of his wish to avoid theorising and stay on the level of description? I doubt if an adequate defence could be mounted in this way, because the limited degree of sociological theorising called for, actually satisfies Wittgenstein's methodological preferences. These preferences are not wholly clear, and may not be completely coherent, but their main thrust lies in two points. I would argue that the self-referential model of an institution meets both. First, Wittgenstein was drawing attention to a simple empirical truth. When we interact with one another, and when we make sense of behaviour, we do not do so by reference to psychological or physiological models of the kind built by scientists (PI: 156–157). This is not a criticism of such models, only of a certain style of speculative commentary on the categories of knowing, intending, feeling, and the like, as they have currency in daily life. To see what these everyday mental concepts really mean, they must be studied as actor's categories. This is not to say that such terms are wholly atheoretical. In a routine way, and in one sense of the word, we 'interpret' the behaviour we see about us. We don't construe it as so much matter in motion, but as meaningful. No one has to teach us to do this. (Here is a point on which Reid's and Wittgenstein's ideas converge, but so do Hume's.) For example, a certain gesture and expression we take to be irritation at an error or the bungled execution of a task (cf. PI: 54). These simple cues are open to view and we respond instinctively, not by speculating about the state of the soul, or the interior machinery of the brain. With regard to the second of these possibilities, Wittgenstein appeared to be in two minds as to whether there was any such machinery to speculate

about. At times he seemed to take it for granted, though deeming it irrelevant, at others he gave free reign to his anti-causal prejudices and appeared to deny it (contrast Z: 304 and Z: 608–610). But this equivocation about the ultimate status of causal, physiological models must be kept completely separate from the empirical truth that using such models is a specialised pursuit, not the currency of everyday understanding. On this question there is no significant divergence between Wittgenstein's insight and the sociologist's concern to understand and analyse actors' categories.

Wittgenstein's second target, when he eschewed theory and explanation, was not the psychologist or physiologist but the philosopher. Here 'explanation' means 'justification' and 'rationalisation', not 'causation'. In this mood, which is never fully differentiated from the former position, the theories he opposed were metaphysical theories. His target was the metaphysical illusion that some deep, justificatory truth could be discerned behind or underneath our practices. A typical passage illustrating this aspect of Wittgenstein's position is the following:

> But how can I explain the nature of a rule to myself? The difficult thing here is not, to dig down to the ground; no, it is to recognise the ground that lies before us as the ground. For the ground keeps on giving us the illusory image of a greater depth, and when we seek to reach this, we keep on finding ourselves on the old level (RFM: VI: 31).

The 'disease' of 'wanting to explain', as Wittgenstein went on to call it, clearly refers here to philosophical not scientific explanation. (And to individualistic speculations, in particular cf. BB: 143). The process of trying to grasp at some illusory basis of rule following, only to find oneself back at the old level was nicely captured in the dialectic of Kripke's presentation. The attempt to answer the sceptic kept bringing us back to mundane things such as feelings and dispositions. It is not, however, the sociologist who is guilty of chasing some illusory metaphysical goal. The collectivist or sociological approach only invokes ordinary human interactions of a kind perfectly open to empirical scrutiny. The victims of illusion are the individualists who, rightly not content with psychological facts, and yet resistant to sociological facts, postulate a special irreducible fact to explain to themselves the nature of rule following.

10

CONCLUSION

The main conclusions for which I have argued are (1) a rule is a social institution, (2) following a rule is participating in a social institution, and (3) an institution can usefully be analysed in terms of collective processes having a self-referring or performative character. The (Wittgensteinian) answer to Kripke's sceptical challenge is that the facts of meaning are the facts of institutional membership. Meaning is a social phenomenon. Obviously, meaning is also a psychological phenomenon. The point is that it cannot be sufficiently characterised in a narrow, individualistic way. The attempt to build an account of meaning on such a narrow basis produces inadequate, or question-begging and obscurantist explanations of normativity. Is this sociological answer to the sceptic a 'straight' or a 'sceptical' solution as those terms are employed by Kripke? Furnishing a sociological fact of meaning brings it under the rubric of a straight answer, and this is how I have presented it. At the same time the answer can, and should, count as a sceptical solution in as far as it acknowledges that there is no individualistic fact of meaning. This, remember, was the real point of the sceptic's challenge. So the present position is both straight and sceptical. It goes along with the sceptic's argument against individualism, but provides a direct answer to the sceptic's wholly reasonable questions about the nature of meaning and rules.

Similar considerations apply to the 'reductive' as to the 'sceptical' character of the enterprise. In one sense the sociological theory is reductive, in another sense it isn't. It could be said to 'reduce' meaning and rules to sociological phenomena. This is how it will appear to advocates of the 'autonomy of meaning' or the so called 'irreducibility thesis' – those who believe, contrary to Wittgenstein, that there is 'some *unique* way' in which future, correct usage is

predetermined by an 'act of meaning' (PI: 188). At the same time the claim has been that institutional and sociological processes cannot be ignored in favour of purely psychological or individualistic ones. This is the anti-reductionist strand in the argument. Confusion can be avoided by remembering that reductionism, like scepticism, should not be thought of as a generalised stance. It is always scepticism directed at such and such a claim, or reductionism with regard to this or that phenomenon. This is perhaps why Kripke, rightly though confusingly, characterised the idea of meaning as a unique, *sui generis* state, as both 'desperate' and, 'taken in an appropriate way', acceptable to Wittgenstein (K: 51). The appropriate way recommended here is that rules and meanings are just as much, and just as little, *sui generis* for Wittgenstein as promises are for Hume.[1]

Hume was my model for explaining rule following in more basic terms. His explanation of normativity in the moral sphere sets out for us the explanatory and logical relations we can expect to encounter in a comparable, sociological account of rules in general. Hume showed how conventions with moral force arise from interactions devoid of an obligatory character, that is, from a pattern of interested and calculative interactions. If we simply move within the completed moral system we can do no more by way of explanation than define one part of the system in terms of another part. This circularity did not, fortunately, suggest to Hume that a more basic explanation was impossible. He saw that the correct response was that the system of discourse as a whole required explanation. Similarly, the self-referential, and hence 'circular', character of discourse about logical compulsion does not preclude explanation directed at the system as a whole. What is required is that we see how non-intentional, non-normative responses can fall into the pattern of what I have called a 'self-referring' system. This can then provide the behavioural and dispositional matrix for the social phenomenon we call 'logical compulsion'. No doubt, ideally, the model should be formulated using terms such as 'feedback loop', rather than 'self-reference', to characterise the initial stages of the process and this, to a great extent, was how Barnes presented the original model in his 1983 paper. Undue sensitivity to terminological issues should not be allowed to obscure the basic point that the question-begging character of discourse within the system does not preclude causal explanation of the system as a whole. If this seems

problematic the answer is: don't think, but look – and see what Hume did.

I began the argument leading to these conclusions by describing Wittgenstein's finitist theory of meaning. Meaning is generated in a step-by-step fashion as we go along. It is not progressively revealed by usage. It does not pre-exist, but is created in response to the sequence of contingencies attending each act of concept application. This is the true significance of the Wittgensteinian slogan that meaning is use. Use is not to be explained by reference to meaning, because use does not come from meaning. Rather, meaning comes from use. Astonishing though it may seem, commentators can still get this the wrong way round. ('Wittgenstein . . . does believe in the future determination of use by meaning . . . ' M: 136 fn 58; cf. 178). The picture of meanings as a pre-existing reality was dismissed by Wittgenstein as myth mongering. His finitism was directed against the tendency to 'reify' meaning. Reification means treating something as an independently existing object or reality when it does not genuinely possess this status, or when it does not have the entity-like character attributed to it. The error creates the temptation to postulate strange, non spatio-temporal entities (PI: 36). Linguistic idealism, or the self-referential and performative model of a social institution, was offered as an alternative account of many of those phenomena, such as meaning, number and rule, which tempt us into reification. One of Wittgenstein's explicit targets in his attack on reified theories of meaning was Frege, but the tendency is a wide-spread one, existing in a variety of guises. An example is to be found in the later work of Popper. We have already encountered Popper, in his earlier work, expressing thoroughly down-to-earth opinions about the social character of scientific method. His orientation is somewhat different in the following comments on the nature of rules and their relation to rule-following behaviour. It is, he said,

> difficult to understand how the physical universe could produce abstract entities such as rules, and then come under the influence of these rules, so that these rules in their turn could exert very palpable effects upon the physical universe.
>
> (Popper 1972: 225)

Popper goes on to describe the problem in terms of 'the influence of the *universe of abstract meanings* upon human behaviour'. The phrase 'universe of abstract meanings', he tells us, is shorthand for

'such diverse things as promises, aims, and various kinds of rules . . . ' (1972: 230–231). Stated in these reified terms one may indeed wonder if the problem of the reciprocal influence of the two 'universes' will prove soluble. The response suggested by Wittgenstein's work is to de-reify the issue. We may do this by seeing Popper's talk about the abstract universe of meanings, rules, and promises as a transfigured and mystified apprehension of the social world. (For a systematic, sociological de-coding of Popper's theory of objectivity see Bloor 1974.) In Wittgenstein's phrase, Popper's language here is 'on holiday' (PI: 38). It needs to be disciplined by being brought back to everyday employment. We must look at the use of the word 'rule' in the context of our customs and institutions. Popper was, however, right to connect rules with promises in his abstract universe, but Hume has shown us how to decode that particular piece of reified talk.

Philosophers of a variety of persuasions share Wittgenstein's sense of the dangers of reification, though they adopt different strategies to overcome it. One widely shared response is the appeal to the concrete realities of neuro-physiology. If meanings do not constitute an abstract universe, perhaps they are states of the brain. If they are not 'out there', perhaps they are 'in here'. One development of this approach is to say rule followers might have something like a computer programme in their heads. Computers can calculate and generate number sequences, so here is a materialistic model of how we can follow the rule for 2, 4, 6, 8, etc. This is a robust approach which has already generated many insights. (See Fodor 1975 and 1987 but also Churchland 1989 and Haugeland 1990.) For our purposes J. J. C. Smart's (1992) discussion, which is directly addressed to Wittgenstein and Kripke, can serve as a representative of this approach. The essence of Smart's argument is simple. The fact of meaning, that is, the fact Kripke's sceptic wanted to identify, can be specified by reference to computer programmes or their analogues in the human brain. What, then, would be the difference between a computer which added and a computer which 'quadded'? It would reside, on this argument, in their inner workings, specifically in the programmes which control them. Now think of the human brain accordingly. The rule which we are really following will be the neural programme controlling the output of responses. That programme determines the meaning of the signs we produce and is the ultimate ground of meaning.

We know from the overall course of our discussion that the

crucial question facing Smart is going to be the explanation of normativity. How does he employ his computer analogy to solve the problem of specifying what counts as a correct employment of a sign in contrast to an incorrect employment? The answer is: he doesn't. And it is impossible to see how he could. A computer with a programme is just a machine with a complicated set of dispositions. These dispositions are, as Kripke rightly said, intrinsically incapable of generating the distinction between right and wrong courses of action. (Unless, that is, we start bringing into the picture the interaction between different computers, and hence begin to model a society.)

Smart disguises the limitations of his model from himself by appeal to the physical integrity of the individual computer. Do we not have an idea of when the machine is working in a way that is normal and proper for it? We can surely differentiate these from cases where it is subject to random, outside interference (e.g. lightning strikes) or other intrusions (e.g. the failure of one of its parts). The computer works correctly when it 'remains an effectively closed system' (Smart 1992: 130). Smart calls this a 'non-normative notion of correctness' (ibid.: 129), because it is, allegedly, a standard belonging to the machine independently of the intentions of the designer or the social norms of the community of users. By now the suspect step will be easy to identify. Smart has transposed old assumptions about the integrity of meanings and concepts into new assumptions about the integrity of mechanical systems. The old reification of meanings is simply replaced by a new reification of the computer itself. The role previously played by the presumed closure of the 'extension' of a concept is now played by the presumed closure of the 'system'. The materialistic character of Smart's approach has been unwittingly compromised. It is materialism hedged around by, and predicated upon, a distinctly non-materialistic set of assumptions about the identity, unity and continuity of a machine. The result is an uncritical use of the very concept that Wittgenstein wanted us to treat with the utmost circumspection: the concept of sameness. It was in anticipation of approaches such as Smart's that Wittgenstein warned against using the image of the machine as a symbol of logical determination and the compulsion of rules (RFM I: 119–122; see also Kripke's discussion in K:35 fn 24). The appeal to the computer fails because, once the machine has been reified into the machine-as-symbol, exactly the same sceptical arguments can be used to make the identity of

the machine problematic as were used to make the identity of symbols, meanings and concepts problematic.[2]

Out of deference for Wittgenstein let me formulate the problem, for the last time, in theological terms. We acquire a set of dispositions: these are the engine driving our rule-following behaviour, and we want to know how they relate to the meaning of that behaviour and the rule follower's intentions. If our dispositions slide into error, in virtue of what do we define this lapse? What was the meaning or logical content of the disposition, as distinct from the purely physical force within it – a force that, it now transpires, inclines in another direction? What can explain this mysterious duality of the disposition? It seems to have become divided into two parts. One part, at the beginning, defined its target. The other part subsequently went off course, lapsed into error, and failed to hit the target. What stops our reasoning back from effect to cause and concluding that the true content is revealed by whatever the disposition comes to? It is as if the disposition possessed not merely its physical attributes but also a non-physical part, the part containing its meaning, as it were the *spirit* or *soul* of the disposition (cf. PI: 36, 530). From a naturalistic standpoint this is unacceptable, though if the word 'irreducible' is substituted for 'spiritual', it apparently still commands respect in some quarters. (For a denunciation of such respect in the case of mathematical rules see RFM V: 16.) Wittgenstein's position was that the mysterious features of our dispositions are just their social aspects, misunderstood and mischaracterised. The non-physical, normative force attending, or informing, the disposition, providing for it a standard of how it should unfold, is not really intrinsic to it. It is imposed on it from without, from society. Understanding the dualism means understanding its social conditions.[3]

The root difficulty in all this is getting the right balance between psychological models, whether provided by dispositions or computers, and a grasp of sociological processes. Hume's work, with its combination of moral psychology and conventionalism, provides a model, though not one without its complexities. We have seen how Hume structured his discussion around the distinction between the natural and the artificial virtues. The natural virtues are innate, psychological phenomena; the artificial virtues are conventional and socially constructed. Should we be saying, then, that there are two kinds of rules, 'natural rules' and 'artificial rules'? This, in effect, is what the individualist says. 'Natural rules' would

be the 'rules' to be found in the state of nature, outside society, on Crusoe's island, along with the natural virtues. When we recall that the natural virtues are only amiable and useful dispositions, then we will see that 'natural rules', too, can be no more than this. They might be remarkable in their complexity and utility, as many animal instincts are. Like these instincts, they may prefigure some of our socially sanctioned performances. The fact remains that this terminology is just another way of invoking the old disposition analysis. It would be a category of 'rules' without normativity.

The best way to use Hume as a model is not to map rules onto to virtues, bifurcating them into the natural and artificial, but to align them with promises. Promise keeping is an artificial virtue and so falls squarely on the side of the conventional. This was the point of the comparison in Chapter 9. It was designed to help us see rule following as an artificial virtue. But couldn't Hume have said there were both 'natural promises' *and* 'artificial promises'? Presumably, in principle, he could. The intention to do X could be glossed as the promise to oneself to do X. Such talk, however, would be obfuscatory. It would be reading back into the individual psyche something whose significance and origin can be best explained by looking at a pattern of interaction between agents. Hume's power to illuminate came from resisting this tendency to obfuscation by keeping his eye firmly on normativity. He had seen what others had not seen: that our capacity to bind ourselves is a deeply problematic phenomenon. We must explain this ability to impose a necessity on ourselves. We must track it back to its origins, not take it for granted, or assume it, or treat it as a postulate. Wittgenstein's insight was of the same kind. This is why it was important to begin our discussion by emphasising the mysterious character of the necessity that attends the following of rules, so we can share the sense that there is a problem here.

When Hume identified promise keeping as a virtue grounded in convention, he took the obvious precaution of defining his terms. He provided a detailed account of the notion of convention and showed in a step-by-step fashion how self-interested actors, thinking strategically, could turn the 'interested affection' against itself (T: 492). He did what Wittgenstein, alas, conspicuously failed to do in the comparable case of rules and institutions. In the course of developing this account Hume exhibited one of the basic logical properties of collective action. The properties of a group cannot be understood as a simple aggregation of the properties of its

constituent individuals. An aggregative model leaves out of account the interactions of the individuals involved. That is to say: it leaves out everything that is characteristically social. In Hume's case the existence of the group can endow individuals with a power they cannot, individually, confer on themselves, namely the power to put themselves under an obligation. He showed how this capacity belongs to the group in virtue of the interactions of its members, the very interactions that constitute it as a group.

Unfortunately, this important logical property of the move from individual to collective action often remains unappreciated. Consider, as an example, the following attack on the sociological or collectivist account of rule following in a 'state of the art' review (Boghossian 1989). The argument is that a collectivist theory of rules will necessarily inherit the shortcomings of an individualist, dispositional account. If the dispositional account fails to explain normativity, then no advantage will be derived from invoking the social group – because what is a group but a collection of individuals? That which, by its nature, and in principle, cannot be accomplished by one disposition, cannot be accomplished by 17,000 dispositions. The author makes his point by reference to a disposition to make a perceptual error, e.g. mistaking a cow for a horse, but the reasoning is the same as for arithmetical calculation. Thus:

> The community . . . however exactly specified, is bound to exhibit precisely the same duality of dispositions that I do. . . . After all, if *I* can be taken in by a deceptively horsey looking cow on a dark night, what is to prevent 17,000 people just like me from being taken in. . . . But, then, any of my dispositions that are in this sense systematically mistaken, are bound to be duplicated at the level of the community.
>
> (Boghossian 1989: 536)

The inference here is similar to Kripke's dismissal of straight sociological answers to the sceptic on the grounds that a community wide version of the individual disposition answer would be open to, 'at least some of the same criticisms' (K: 111). But while Kripke's inference is guarded and implicit, Boghossian's is more forthright, and clearly fallacious. It is not true that a disposition to error on the part of every member of a group is bound to be duplicated at the level of the community, so that the community *qua* community will be so disposed. The logical route from 1 to 17,000 is not to be travelled by simple addition. It is certainly possible for a disposition to

be duplicated at the collective level in this way, but it is not a necessity. There is no 'bound to be' about it: everything depends on the contingencies of interaction. Popper's philosophy of science provides a celebrated counter-example to the alleged necessity. Suppose, counter-factually, that every member of the scientific community lacked the power of self-criticism. Would that imply that science itself, the collective enterprise, would be unself-critical? Not at all. Science could be a self-critical enterprise even if every single scientist lacked this property. The trick would be to get them to criticise one another. Engineering an individualistic, competitive social structure would be one way to achieve this.[4]

That fallacies of this kind still have currency shows the level of sociological awareness being brought to bear on the issue of rule following. No wonder, when Wittgenstein failed to tell us what he understood by an institution, a yawning gap was left in the secondary literature, a gap that is apparently invisible to otherwise well-informed commentators who believe that everything that could be said in response to Wittgenstein has been said. Nothing could express more eloquently the individualist tenor of contemporary philosophy. It is a sign of the times.

The argument of the book is now finished, but I should like to add something of a more personal kind about the purposes behind it and the goals directing it. In 1973 I published a paper called, 'Wittgenstein and Mannheim on the Sociology of Mathematics' (Bloor 1973). The argument of the paper was that, in principle, Wittgenstein had solved for us one of the outstanding problems in the sociology of knowledge. He had revealed the social character of rule following. In doing so he removed a stumbling block that had brought even so determined an advocate of the sociology of knowledge as Karl Mannheim to a stop. Mannheim had conceded that the sociologist could locate sociological causes (or what he called 'existential determinants') only in those cases where a line of thought or behaviour *deviated* from that indicated by the meanings of the theories, concepts, ideas or rules guiding an historical agent. As long as an agent was determined by the unfolding content of his ideas and beliefs then, Mannheim implied, existential determinants were not at work. The sociologist might, in principle, explain how an agent came to have certain beliefs or commitments in the first place, but not how the agent then acted under their guidance. At that point logic and rationality themselves take over as explanatory principles. It is as if Mannheim said to himself: as long as we are

logical then logic guides us, but if we are illogical then, and only then, does society intervene. In short, for Mannheim, society distorts the logic and rationality inherent in our concepts. The sociology of knowledge was therefore impossible, except as a study of generalised, predisposing conditions. We must settle for a sociology of error. Mannheim didn't express himself in quite this way, but this was the import of what he said when he talked of the 'inner dialectic' of concepts (Mannheim 1936).

My claim was that in his discussion of rules, Wittgenstein had put his finger on precisely the issue that vexed Mannheim. Following a rule epitomises the case Mannheim could not understand, the case preventing him from formulating a fully general sociology of knowledge. I argued that Wittgenstein did not merely identify Mannheim's problem with great clarity, he solved it as well. His solution was to see rules, and the so-called unfolding of meaning, as themselves social processes.

It was only a mythical conception of these things that caused sociologists to limit their legitimate aspirations. In Wittgensteinian terms, Mannheim and those numerous philosophers who still think he was right, were simply bewitched by language. They were the victims of philosophical confusion, and in the grip of a false theory of meaning, namely, meaning determinism. The idea that concepts have an inner dialectic is just another way of expressing this doctrine. A proper Wittgensteinian analysis of rule following, as part of a shared language game, and the analysis of rules themselves as institutions, would liberate sociologists of knowledge from their self-imposed fetters, and encourage them to address the real but difficult problems of their discipline.

For me this has always been the primary interest of Wittgenstein's work: its power and utility as a resource for the sociology of knowledge. Elsewhere I tried to fill out the picture of Wittgenstein as a sociologist of knowledge (Bloor 1983). But, as far as my own contributions and understanding were concerned, I left the central issue of rule following, the nub of the whole problem, standing where it had stood in the Wittgenstein and Mannheim paper. Unfortunately, in that paper, the Wittgensteinian solution to Mannheim's problem was only presented in generalised terms, and looking back I am conscious of how schematic it was. The concepts Wittgenstein used for redescribing rules as social phenomena were correctly identified and labelled. Custom, convention, norm and institution were to provide the apparatus of analysis and

explanation. Logical compulsion was to be seen as a species of moral compulsion. In its essentials the argument didn't go much beyond this. These ideas themselves were not filled out or given substance. The paper showed, or was meant to show, that a solution to Mannheim's problem existed, that its existence could be relied upon, and was to be found in a certain specified terrain, but it didn't produce the object for inspection.

I now see why the first attempt to present Wittgenstein as a sociologist of knowledge was limited to a sketch of this character. Despite his insights and sociological tenacity, Wittgenstein's own work itself exhibited, at certain crucial points, that same schematic form. My derivative argument merely inherited and transmitted the shortcoming. This book is an effort to make good that deficit and go some way towards exhibiting the sociological substance of rules and rule following. That is why I have now had to supplement and continue the Wittgensteinian analysis by using the self-referential model of an institution developed by Anscombe and Barnes. This, too, can be seen as relatively schematic and in need of further development, but it represents a step forward of the kind needed. The next step must be to produce richer theoretical models of this kind and also to identify these self-referential processes in historical case studies. As examples of such work, I would point to Barry Barnes' *The Nature of Power* (1988) and Steven Shapin's *A Social History of Truth* (1994). The time has come, then, not merely to learn from Wittgenstein – once his work has been brought back into proper focus – but to extend, deepen and, where appropriate, modify his style of analysis. The slogan must be: to understand Wittgenstein is to go beyond him.[5]

NOTES

PREFACE

1 In a useful article Champlin (1992) distinguishes between 'his custom' and 'the custom', the first being an individual, the second a social, phenomenon. He argues (1992: 292) that Wittgenstein's use of the word '*Gebräuche*' in *Philosophical Investigations* (PI: 199) is a significant indicator that the social rather than the individual connotations of 'custom' were intended. I find Champlin's argument convincing, though my own reading will depend on the overall architecture of Wittgenstein's position rather than linguistic considerations.

2 Support for an individualist account of rule following in general, and sometimes for McGinn in particular, and/or criticisms of Kripke can be found in Arrington (1991), Baker and Hacker (1984), Blackburn (1984), Boghossian (1989), Budd (1984), Carruthers (1986), Coates (1986), Collins (1992), Gilbert (1983), Ginet (1992), Goldfarb (1985), Harrison (1991), Hoffman (1985), Loar (1985), Luntley (1991), McDonough (1991), McDowell (1984), (1991) and (1992), McKinlay (1991), Shanker (1984), Smart (1992), Werhane (1987), Wright (1986). Recent statements of support for collectivism in general, and/or Kripke, can be found in Armstrong (1984), Champlin (1992), Malcolm (1989), Margalit (1992), Peacocke (1981), Schiffer (1986), Summerfield (1990), Williams (1983), (1985), (1990), (1991), (1994a) and (1994b). The last mentioned sequence of papers, by Meredith Williams, represent a powerful and sustained attack on individualism.

1 INTRODUCTION

1 The practical person's impatience with reflections on what is normally taken for granted is not, of course, identical to the philosopher's defence of common sense. The so-called 'relaxed common sense' of the philosophers is actually far from relaxed: it is sophisticated, complex, and strained. The point of putting them together is that both have the effect of blocking empirical enquiry into, and constructive theorising about, common sense practices. I am aware this is a harsh judgement and that

there is an argument to be had about what constitutes blocking empirical enquiry. For the moment I leave these as assertions. Later, in Chapter 7, I shall provide some examples of what I mean by this. I am also aware that many defenders of common sense take their lead from statements made by Wittgenstein about the proper methods of philosophy. He explicitly said the aim was not to construct theories. His precepts, however, were not wholly at one with his practice. It is what he did rather than what he said which will be my concern. That his precepts may have inhibited his practice, and that his constructive work was hesitant and incomplete, I do not deny. In selecting which parts of Wittgenstein to cherish, and which parts to regret and discard, we reveal our intellectual values. I have spelled out the basis on which I select what I take to be the important doctrines in my 1983 book. I also make clear which aspects of his thinking I am committed to contradicting. These are, roughly, his meta-level assertions about philosophical method.

2 Is meaning determinism a doctrine about our grasp of meanings or about meanings themselves? Does it tell us about what causes our behaviour or about the logical standards by which it is to be assessed? The answer is that it deals with, or runs together, both sets of themes. We act as we do because we grasp meanings, and meanings (by their very nature) imply or constitute standards of right and wrong usage. It is, however, difficult to give a sharply defined statement of meaning determinism. I have never encountered a clear formulation, and never expect to. This is because meaning determinism represents a reified version of phenomena whose true location remains unacknowledged. It is also a doctrine more often implied than stated. It lives in the realm of tacit assumptions, intuitions and metaphors. For example, the process called 'grasping' a meaning has never been made clear, though it is sometimes said that the process is not a 'causal' one. The claim appears to be that 'efficient' causes are not in question, though 'final' causes may be. Meaning determinism often makes sense as a tacit form of teleology. See Bloor 1973.

3 As an example of the problem confronting individualists when trying to account for normativity I would point to Baker and Hacker's book *Scepticism, Rules and Language* (1984). Instead of being given a clear account of normativity the reader finds that the concept is smuggled into the discussion via references to '*regularities* of action of sufficient *complexity* to yield normativity' (Baker and Hacker 1984: 42) and the unexplicated concept of 'normative regularity' (ibid.: 13, 14, 16, 19, 21). How normativity is generated from regularity and complexity is never explained. Further examples of this failing will be given later. For a succinct but effective criticism of Baker and Hacker see Sartorelli's (1991) review.

4 The name (and this writer's awareness of the importance of these themes) comes from Hesse's important (1974) work. Hesse's network model of language is strongly influenced by Wittgenstein's work and explores the consequences of his approach for scientific concepts and the confirmation of scientific laws and theories. Hesse relates these ideas

to Carnap's 'instance confirmation', where the measure of confirmation refers to the next instance of a putative law, not the law itself, cf. p. 167. For a valuable account of meaning finitism, also influenced by Hesse, see Barnes (1982) especially Chapter 2. The themes to be developed under the heading of 'finitism' are by no means absent from the philosophical literature devoted to Wittgenstein and rule following, but they are certainly under-represented. The nearest I have come across is Wright (1986) who calls the position 'constructivism'. One of the sources of Wright's reading is Dummett (1959) who represents Wittgenstein as a radical constructivist. I have discussed Dummett's account of Wittgenstein's mathematical views in Bloor (1983).

5 For a criticism of normative determinism, taking the sociologist Talcott Parsons as the target, see Barnes 1988: Chapter 2, Section 2.1.

2 MEANING FINITISM

1 I am reading the passage quoted above, from RFM I: 3, as much 'ontologically' as 'epistemologically', but it has been suggested to me that 'strictly' it is only epistemological in its import. Is it not about how we know what to say after 2004, as distinct from what makes it correct that 2006 is the answer? I think this is a one-sided rendering of what is said in the passage. Formulations of meaning determinism typically run together these two issues, and so did Wittgenstein's rebuttal. Remember that, for meaning determinists, knowing the correct step arises from 'grasping' that which makes the step correct, that is, the pre-existing meaning of the rule. So the two questions of how it is known, and what makes it so, are asked and answered in ways that make them dependent on one another. Correspondingly the rejection of this doctrine addresses both aspects. Thus Wittgenstein was saying that the automatic character of our responses (sustained by consensus) dispensed with the 'grasping' process, and hence with the pre-existing reality of the rule – the reality which, for the meaning determinist, settled the correct answer in advance. This is the assumption which becomes superfluous.

2 Meaning finitism amounts to more than the simple claim that words can change their meaning, as when we observe that the word 'specious' meant something different in the eighteenth century from what it does now. Such a phenomenon can easily be accounted for by a meaning determinist – it is just a case of switching from one rule for using a word to another, where 'rule' is thought of according to meaning determinist principles. Meaning finitists would want to add the assertion that there is indeterminacy within each of these rules.

3 Calling Wittgenstein a meaning finitist may seem to conflict with his known rejection of 'finitism' and 'intuitionism' in mathematics itself. These doctrines represented an attempt to limit the proof procedures that were accepted as legitimate. They were informed by the idea of a 'crisis' in the 'foundations' of mathematics and the need for a change in mathematical practice – ideas that were foreign to Wittgenstein. Meaning finitism has affinities to such doctrines, but the two are not the

same. It is possible to be a meaning finitist without being a mathematical finitist in the above sense. Meaning finitists need have no quarrel with mathematical practice, as long as those practices are analysed finitistically. This might involve rejecting the philosophical gloss put on them, but not the practices themselves. The model to keep in mind is the example used in the text: that of giving a finitistic account of concepts of infinity – which did not involve rejecting talk of infinity as meaningless or illegitimate.

4 Baker and Hacker (1984) proceed as if we can 'grasp' the extension of a rule. On their view a rule and its extension, that is the entire class of its correct applications, are linked together by what they call an 'internal relation'. To understand rule R means knowing *these* are its instances. Accepting other, and different, instances or applications would imply that some other rule than R, or no rule at all, was being followed. To follow a rule we allegedly 'grasp' the internal relation connecting it with its application. 'To understand a rule is to grasp an internal relation between the rule and its (potential) extension' (Baker and Hacker 1984: 76). 'For what "understanding" here consists in is a grasp of internal relations between rules and their applications' (ibid.: 93). Williams (1991) offers well-aimed criticism of this aspect of Baker and Hacker's argument, concluding that it results in 'the mystification of grammar' (Williams 1991: 103).

5 Wittgenstein is not alone in his rejection of the standard apparatus of philosophical semantics. An increasingly influential line of criticism has come from philosophers impressed by computer models of the brain based on neural networks. Neural networks do not work by manipulating symbols according to rules. They have an entirely different architecture based on parallel, distributed processes. Knowledge is not embodied in sentential form, but as a distribution of weightings which determine the spread of excitation within a network. On this view the individual would only possess a diffuse, holistic disposition to act. There are no 'rules' in the head, or verbal formulae in the so called 'language of thought'. It is tempting to suggest that if there were a sophisticated, technological embodiment of Wittgenstein's account of thinking, then it would be here. Of course, the distinctly scientistic ideology of many PDP enthusiasts is at odds with Wittgenstein's anti-scientific temper. A good introduction to the non-propositional approach is Churchland (1989).

3 RULES AS INSTITUTIONS

1 The Humean definition has been rendered into formal terms by Lewis in his book *Convention* (1969). For Lewis, conventions are stable solutions to co-ordination problems. His 'first, rough definition' of a convention is as follows. A regularity R in the behaviour of members of a population P, when they are agents in a recurring situation S, is a convention if, and only if, in any instance of S among members of P, (1) everyone conforms to R, (2) everyone expects everyone else to conform to R, (3)

everyone prefers to conform to R on condition that others do (Lewis 1969: 42). Later refinements include relaxing the condition that everyone conforms to R.

2 There is no unique way of responding to Wittgenstein's failure to provide an account of institutions. (Just as scientific theories are under-determined by experimental results, so the interpretation of a text is under-determined by the words on – or not on – the page.) One possibility, and this seems to be the standard response, is to reason that if Wittgenstein did not tell us what he meant by an institution then he did not intend to tell us. So it must have been intrinsic to his enterprise not to engage in any such explicit or constructive analysis. On this reading, if we want to understand Wittgenstein we must create a sense of what he was about which would render the absence of a theory of institutions natural and proper rather than shocking. Perhaps Wittgenstein was simply describing common-sense structures of thought from, as it were, the 'inside'. His philosophy is then merely a reminder of our common-sense understanding and its only role is the negative one of removing the myths that may accumulate around that understanding. Such an enterprise would not call on any higher level, or more generalised form, of understanding than those already embedded in our common-sense grasp of our culture. We do not need a novel theoretical or analytical standpoint; the social agent's or social actor's own categories will suffice. This reading certainly fits some of Wittgenstein's expressed sentiments. There are clear ways of attaching it to points in the text. It also has another conspicuous virtue: it is a 'charitable' reading. It refrains from identifying error in the text, interpreting the features in question as rational expressions of the author's intentions, while the understanding of that intention is adjusted accordingly. Charity, along with faith and hope, are indeed great virtues, but they are not the only ones, and sometimes they can be a little bland, and even dull. This charitable reading has found ample expression in the literature and it is time to assess its results. This takes us into the realm of judgement, and such judgements cannot be decisively justified. I shall therefore allow myself to be dogmatic and assert that 'charitable' readings, along the lines just sketched, seem to me to have led to the trivialisation of Wittgenstein's work. We should therefore be ready to explore alternatives. Even if this opinion were to be rejected, and current Wittgensteinian philosophy were seen as generating a wealth of fascinating insights, it would still be proper to develop alternative readings. Only when such alternatives are on the table can a properly comparative judgement of the virtues of the different readings be attempted.

3 There are interesting analogies between, say, being a coin and being a window or a tile. A window is not just a hole in a surface, it is something made and used and thought of in terms of its utility for human habitation. The same applies to tiles. Not any piece of slate is a tile. Their significance is not, as it were, wholly contained within the limits of the object, but depends on how they are treated and regarded. These points do not destroy the contrast I am drawing but, rather, they indicate ways

in which a sociological perspective can be generalised. Anscombe (1976) called such cases 'partial idealism'. For a fuller discussion of this continuity see Barnes (1988).

4 To some extent this argument coincides with that of philosophers who have maintained that social facts are irreducible. For a classic statement see M. Mandelbaum's 'Societal Facts' (1955).

5 Barnes provides the more systematic formulation of the two. His work is also clear in acknowledging the naturalistic basis of the analysis. Anscombe by contrast explicitly wants to harness the argument to an anti-naturalistic enterprise (Anscombe 1978b, reprinted 1981: 97). Ideas similar to Barnes' have also been developed recently in Searle's (1995) *The Construction of Social Reality*. The difference between Searle's and Barnes' earlier formulation lies in (a) Searle's rejection of finitism, and (b) his individualism. These differences are identified and discussed in Bloor (1996a). A version of linguistic idealism can be found running through Winch's *The Idea of a Social Science* (1958), e.g. 'social relations really exist only in and through the ideas which are current in society' (Winch 1958: 133, see also 118, 123 and 131). Winch shares Anscombe's anti-naturalist tendency.

6 The notion of an 'actor's category' is not itself an actor's category. It is part of a perspective which is beginning to go beyond that of actors themselves and represents a new level of self-awareness. Indeed, the whole analysis of institutions offered here is part of such a project, even though it is central to that project to appreciate that rules, and other institutions, are themselves intrinsic to the original level of self-understanding, that is, they are, and are constituted by, actors' categories.

7 I am thinking here of cases like, 'Why must I move the king like this on a chessboard?', rather than, 'Why play games?' While some wholesale and pragmatic answer like 'for relaxation' might suffice for the latter question, the former kind merely lead us around within the game.

8 The notion of a natural kind is here being used uncritically in order to draw a contrast with social kinds. Nothing would be gained if the complexities of our (finitistic) use of natural kind terms were brought into the story. It is perhaps appropriate to emphasise that finitism is in no way at odds with the use of natural kind terms. It would be a misunderstanding to see finitism as rendering the idea of a natural kind untenable – as an actor's category. For example, suppose it were said that the idea of a natural kind rests on nature setting the standards of what counts as similar to what, where these standards are wholly independent of what we happen to find it natural to classify as similar to what. That idea is, indeed, incompatible with finitism. Fortunately, for finitism, it is an incoherent idea, and a strangely anthropomorphic one at that. We set standards, nature can do no such thing.

4 CONSCIENTIOUSNESS

1 One feature of the religious experience of some Protestant sects is the phenomenon of glossolalia, or 'speaking in tongues'. Those moved by

this experience take themselves to be speaking in a foreign and unknown language. I understand that to the unbeliever it sounds like inarticulate noise, or nonsense vocalisation. I have often wondered whether this strange identification of the inarticulate with the authentic provides a further point of contact with Wittgenstein's private language discussion. Recall that he said, 'And sounds which no one else understands but which I '*appear to understand*' might be called a 'private language' (PI: 269), cf. also the reference to speaking in tongues in PI: 528. A remark in Malcolm's *Memoir* suggests these connections may not be wholly fanciful, 'The *Journal* of George Fox, the English Quaker, he read with admiration – and presented me with a copy of it' (Malcolm 1958: 71–72).

2 Wittgenstein himself almost gave the game away. He acknowledged the possibility of a social form involving the institutionalised acceptance of something akin to a private language, that is to say, a social form celebrating an extreme individualism. In the section of the *Investigations* dealing with inspiration, the inner voice, and the like, he considered the possibility of a form of arithmetic which was inspirational, where calculating would become more like composing. He did not, however, pursue these hints. Had he done so it would have helped to expose some of the logical limitations of the anti-private language argument as a weapon against extreme cognitive individualism. For a fuller discussion of this theme see Bloor (1997) forthcoming.

3 A comparable failure is exemplified in the following remark directed at Kripke: 'It is hard to see any content in a general notion of "role and utility in our lives": what does it rule out?' (Goldfarb 1985: 481). Another example is provided by McDowell (1981: 156). In discussing how we generate the 2, 4, 6, 8 sequence he rules out (1) Platonism and (2) psychology ('a catalogue of how humans act and feel when they engage in deductive reasoning'). We need, he says, an 'intermediate' position. This leaves him poised to offer a sociological approach, but no such thing is forthcoming. Instead he opts for something identified by non-social categories and selects a quotation from Wittgenstein referring diffusely to an unspecified 'human natural history' (RFM VI: 49). All Wittgenstein's talk of customs, conventions and institutions passes by unremarked.

5 RULE SCEPTICISM

1 Kripke says Jones is entitled, subject to correction from others, to follow his confident feelings about the answer to an addition problem (K: 90). This is Kripke's way of talking about the instinctive component in the rule-following process which allows us to go forward without always looking over our shoulder. Baker and Hacker (1984) fasten on this passage, clearly believing they have caught him out. They say Kripke's position is 'bizarre'. 'That one feels confident is not what *entitles* one to say that one means W by "W". And that one is *inclined* to answer

thus-and-so is not what *entitles* one to judge one's answer to be correct' (Baker and Hacker 1984: 34). But Kripke didn't say Jones' inclinations *gave* him any entitlement. He said Jones *has* an entitlement *to* his inclinations. The entitlement doesn't come from the inclinations; it has those inclinations as its object. The origin and cause of the entitlement, or right, is the group practice. It is, indeed, the whole point of the theory Kripke is attributing to Wittgenstein that an individual's inclinations or dispositions can't generate normativity.

2 An attempt to overcome this part of the argument is to be found in Stabler (1987). He says a finite physical system can embody an infinite function. We just need to think of it continuing on in the same way. We need a concept of its normal operation, i.e. what it would do if all interfering conditions and limitations were removed. Everything hinges on conditional propositions about the system and the identification of the physical laws involved in its normal operation. Can't these be fixed by the vocabulary of the physical sciences, asks Stabler? The answer appears to be positive, and to this extent he makes it plausible that we can extrapolate the idea of a disposition. Stabler, however, takes no account of the finitist character of the concepts used in this specification or of their ultimate dependence on processes of ostension. The sceptic would simply repeat his challenge by applying it to the specification of normal functioning, or to whatever other processes of identification were used. Other versions of this attempt to give a straight answer to the sceptic by refining the appeal to dispositions are to be found in Coates (1986), Forbes (1984), Pettit (1990) and Smart (1992). Smart's argument will be further discussed in the concluding chapter.

3 See RFM VI: 44: 'But aren't we guided by the rule? And how can it guide us, when its expression can after all be interpreted by us both thus and otherwise? I.e. when after all various regularities correspond to it. Well, we are inclined to say that an expression of the rule guides us, i.e. we are inclined to use this metaphor.' I have counted some forty passages in the RFM alone which carry a similar message and invite an analysis in terms of assertion conditions. Amongst the more striking see RFM I: 110, 118, 130, 156; and I: Appendix III, Section 8.

4 It may be said that this fails to meet the problem. Isn't the sociological view that there is no 'natural fact of the matter' that two things are similar, so that every judgement of similarity must be based on a communal agreement that they are similar? But then that agreement itself must be subject to further agreement (that it *is* an agreement) – and so a regress takes hold. Saying we all go on automatically, so the objection goes, misses the point. What makes it the case that the ways individuals automatically go on count as their being in agreement? There is still the need for agreement that it is an agreement. But is there? The objection under-estimates the force and scope of Wittgenstein's point by smuggling in the idea that social processes are like rules for following rules. The basis of the mistake lies in representing sociological finitism as saying there is 'no natural fact of the matter'. What is actually implied by this doctrine is that there is no predetermined and

uniquely correct description of any matter of fact. It does not say that there is no independent reality or no real dispositions which may be contingently aligned in such a way that social interactions can be built up. The entire sociological account rests on given, automatic dispositions – recall what Wittgenstein said about ostensive training. Whilst these may always be overridden, they are always there to fall back on. If they were not there, social life would be impossible, including agreements and disagreements.

6 THE ANALOGY WITH VON MISES

1 I cannot judge the accuracy of Kripke's impression that von Mises' argument is almost universally rejected, but there can be no doubt that it has supporters amongst free market enthusiasts. For an example of how the situation looks from the standpoint of a supporter of von Mises, see Steele (1992). One of von Mises' explicit targets was the Vienna circle's Otto Neurath, a socialist who wrote about the transformation of capitalism into various forms of centralised non-market economy (see Neurath 1973: Chapter 5). Neurath acted as an economic adviser to the short-lived communist government in Munich immediately after the First World War.

7 INDIVIDUALISM

1 Here is an alternative reading of paragraph 197 that has been suggested to me. This places a different emphasis on the words 'and yet isn't present'. It has Wittgenstein saying that there is nothing wrong with thinking that future use is indeed present to the mind. All that Wittgenstein was telling us to avoid, on this reading, is thinking simultaneously that future use is, and is not, present to the mind. In other words, Wittgenstein was simply enjoining us not to think a contradictory thought. I reject this reading as over-ingenious. First, I know of no philosopher who has ever embraced this contradiction, so there is no target or motive for Wittgenstein saying this. Second, this reading is retrograde because it opens the door to more or less disguised versions of the very doctrines he was opposing – as the cumulative effect of the rest of his argument ought to make clear.

8 ISOLATION AND INNOVATION

1 An important consideration in the argument is that it applies to a new rule R, where the judgement of novelty resides with the rule-following community itself. Novelty here is an actor's category. This is why Wittgenstein can rightly say: 'It is possible for me to invent a card-game today, which however never gets played. But it means nothing to say: in the history of mankind just once was a game invented, and that game

was never played by anyone' (RFM VI: 43). The new card game mentioned in this passage will be a variant of existing games and, I take it, sufficiently close as not to count as a genuine innovation in the sense under discussion. In the terms of the self-referential model the point is this: the boundaries of the institution, what counts as part of it, or what lies outside it, is constitutive of the institution itself, and sustained by the same mechanisms.

2 In a later publication McGinn has adopted a position very similar to Millikan's. His new formulation of naturalism will thus inherit both the strengths and weaknesses of Millikan's. See McGinn (1989: 143).

3 Millikan herself has not always been so sure about the character of the normativity that can be grounded in evolutionary and functional considerations. In her impressive *Language, Thought, and other Biological Categories* (1984), functional considerations are called 'quasi-normative' (ibid.: 5) or even 'non-normative' (ibid.: 17). These qualifications have disappeared in the 1990 response to Kripke's sceptical challenge. I think the earlier, more cautious formulation is the right one. There have been other attempts to extract the normative character of meaning from (non-social) nature, e.g. Fodor's refinement of the causal theory of perception and thought in which an inner state or token means X in virtue of being caused by X. He seeks to explain error, that is an X-meaning token that is caused by not-X, in terms of an asymmetrical dependence relation, i.e. the parasitic character of error on truth (cf. Fodor 1987: 106). As with Millikan's argument it is unclear how these considerations are to operate as actors' categories without bringing in social processes. For a perceptive discussion see 'Social Norms and Narrow Content' (Williams 1990).

4 To see how historians of biology handle real episodes of this kind see e.g. Desmond (1989), Rudwick (1972) and Winsor (1976).

5 Wittgenstein discussed both the Zermelo and Cantor cases. It is clear that he wanted to convey the element of choice available to the community of mathematicians in deciding the meaning to be put on the procedures these innovators had made available (cf. RFM II: 22; V: 7). His account of Cantor's diagonal procedure also brought out the performative character of these decisions. '"So it produces a series that is different from all of these". Is that right? – Yes; if, that is, you want to apply these words to the described case' (RFM II: 8; cf. RFM IV: 35).

6 In his 'The Historical Structure of Scientific Discovery' (1962), Kuhn distinguishes two kinds of discovery (1) within-paradigm, or normal science discovery, and (2) paradigm-forming discovery. The former may be attributable to an individual acting at a particular time and place. The latter involves a shift in the community's understanding of the world and the scientist's actions in it. These are not point events. Here we have the full process model of discovery. A similar dichotomy had earlier been developed by Toulmin who distinguished between what he called 'natural history' discoveries and truly 'scientific' discoveries. See Toulmin (1965). The dichotomy adopted in Kuhn and Toulmin's early formulation is undermined by the finitist theory of meaning. There are

now no pre-existing frameworks, such frameworks have to be manufactured as the process unfolds. There is really only one kind of discovery as far as underlying semantic principles are concerned, namely the second, more complex type. In this respect the difference is one of degree. Normal science is truly creative too – as Kuhn always stressed. Such qualitative differences as there are, along other dimensions of the situation, derive from the actor's perception of the continuities or discontinuities within their own practice. These decisions about the sameness of their practice are themselves to be understood as institutions, that is, as constitutive of the sameness or difference.

7 On Priestley see Kuhn (1962); on Mendel see Brannigan (1981) and Olby (1979); on Boyle see Shapin and Schaffer (1985).

9 RULES AND THE STATE OF NATURE

1 Reid's internal relation argument has found supporters in Carritt and Searle. It is discussed and effectively rebutted in J. L. Mackie's sociologically oriented *Hume's Moral Theory* (1980: 98). The attempt to generate obligation out of meaning, which is the underlying purpose of the internal relation argument, is just another variation on meaning determinism. As I have indicated in note 3 to Chapter 1, a version of the internal relation argument plays a central role in Baker and Hacker's (1984) individualistic reading of Wittgenstein. They say collectivists treat the relation between a rule and its application as an external, rather than an internal, relation, that is, as a contingent, causal link. Building consensus into the definition of correct rule following, they say, means abandoning the insight that rule and application are internally related. In reply, I would claim that the circularities to which Baker and Hacker are responding find their most natural and revealing explanation in the context of a consensual, self-referring model of rules as institutions. For a further discussion of the use of the internal relation argument by Baker and Hacker see Bloor (1989) and (1992).

2 See Reid's (1785) *Essays on the Intellectual Powers of Man*, Essay IV, Chapter I, Sect. xxiii. Modern heirs to the Reidian tradition of common sense, intuitionism, can be found amongst phenomenologists. It is interesting to notice that while McGinn effectively wants to place Wittgenstein in this tradition, self-professed adherents to that approach are highly aware of the differences, and criticise Wittgenstein accordingly. Thus J. N. Findlay, writing from a 'traditional acts-of-mind' (1984: 191) standpoint criticises Wittgenstein for *not* adhering to the line McGinn says he *does* adhere to.

3 Could it be argued that Wittgenstein's opposition to moral foundations implies a rejection of a sociological enterprise such as Hume's? Was not Hume trying to provide a sociological foundation for morality? If so it would cut across the sentiment in the criticism of Schlick just given. The point could be generalised to the conclusion that the sociological approach advocated in response to Kripke's sceptic is likewise an

attempt to provide a sociological foundation for rule following, and would equally clash with Wittgensteinian principles. Such an argument, however, depends on equivocating with the notion of 'foundation'. Sociologists – like Hume – are concerned with explanation, while the traditional foundationalist enterprise was concerned with justification. Wittgenstein's opposition was to foundations which glow with the light of their self-evidence and carry with them some manner of self-justification. (There is an analogy here with his rejection of the interpretation model of rule following which has it resting on a last interpretive step which had, of course, to be mysteriously self-interpreting.) Hume was certainly not providing a foundation for morality in that sense. Indeed he was criticising Reid's attempt to furnish a justificatory, rather than a truly explanatory foundation. The corresponding point applies to the sociologist who provides an explanatory, sociological 'foundation' for rule following.

4 Baker and Hacker (1984) quote this passage. 'In the light of Wittgenstein's opposition to rationalistic modes of thought it is interesting to see how these individualistic and rationalistic commentators gloss the ideas in it. They use it in the course of rejecting the idea that Wittgenstein was a sceptic. It would be surprising, they say, to find Wittgenstein, "accepting the sceptic's premises, the 'doubts at bedrock', rather than showing that they are "rubbish"' (1984: 5). But 'showing' (that is, in some way demonstrating or establishing) that the sceptic's doubts are rubbish was precisely the thing he said we could not do. We can, with his blessing, call them rubbish; but show it, and justify our dismissal as a piece of knowledge, we can't. Baker and Hacker have inverted the meaning of this passage in the course of casting it into their own words. They have turned Wittgenstein from an anti-rationalist into a rationalist.

10 CONCLUSION

1 Stated in general terms the relation between processes at the social level and at the individual level may be characterised as one of 'weak supervenience'. The label is due to Haugeland, and is meant to capture a general or ultimate dependence, but one which cannot be expressed in terms of token-by-token identity. Events specified at one level cannot, necessarily, be matched with events specified at the other. He illustrates this by reference to the relation between descriptions of the behaviour of superimposed water waves and descriptions of the behaviour of water molecules (Haugeland 1982). For similar comparisons concerning the relation of macro and micro properties of physical systems see Barnes' remarks on structural explanations in his *Interests and the Growth of Knowledge* (1977). Language games can also be thought of as superimposed, see Bloor (1983: 110).

2 A further striking example of reifying dispositions is provided by Ginet (1992). In a paper subtitled 'answering Kripke's objections', i.e. Kripke's objections to the disposition theory of meaning and rule following,

Ginet effectively inverts the issue he claims to be confronting. He glosses the concept of disposition in such a way that it takes on all the (mysterious) properties of an (uncritical) appeal to meaning. Instead of finding the notion of meaning explained by reference to dispositions, the reader suddenly realises that dispositions are being explained by appeal to meanings. The rule follower, it transpires, acquires the relevant dispositions from having understood the rule: 'the subject has the dispositions because she grasps that rule' (Ginet 1992: 60). That this should be put forward as an 'answer' to Kripke is, to say the least, extraordinary.

3 In his discussion of these themes Wittgenstein can be seen as an unwitting, but impressively subtle, member of the Durkheimean tradition in the sociology of knowledge. See Durkheim's discussion of the soul in his *Elementary Forms of the Religious Life* (1915) and his 'The Dualism of Human Nature and its Social Conditions' (1914).

4 The dangers of inferring the properties of groups from those of their constituent individuals have been well known to social scientists at least since Olson's *The Logic of Collective Action* (1965).

5 I fear that I have written my two books on Wittgenstein in the wrong order. This book could be looked upon as preparing the ground for the further work of construction and bridge-building of *Wittgenstein: A Social Theory of Knowledge,* Bloor (1983).

BIBLIOGRAPHY

Anscombe, G. E. M. (1969) 'On promising and its justice, and whether it need be respected *in foro interno*', *Critica* vol. 3 (reprinted in *Collected Papers*, 1981, vol. III: Chapter 2).

—— (1976) 'The question of linguistic idealism', *Acta Philosophica Fennica* vol. XXVIII: 188–215 (reprinted in *Collected Papers*, 1981, vol. I: Chapter 13).

—— (1978a) 'On the source of authority of the state', *Ratio*, vol. 20: 1–28 (reprinted in *Collected Papers*, 1981, vol. III: Chapter 14).

—— (1978b) 'Rules, rights and promises', *Midwest Studies in Philosophy* vol. 3: 318–323 (reprinted in *Collected Papers*, 1981, vol. III: Chapter 10).

—— (1981) *The Collected Philosophical Papers of G. E. M. Anscombe* vols I–III, Oxford: Blackwell.

Armstrong, B. F. (1984) 'Wittgenstein on private language: it takes two to talk', *Philosophical Investigations*, vol. 7: 46–62.

Arrington, R. L. (1991) 'Sign-post scepticism', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 13–33.

Austin, J. L. (1961) 'Performative utterances', *Philosophical Papers*, Oxford: Clarendon: Chapter 10.

Ayer, A. J. (1954) 'Can there be a private language?' *Proceedings of the Aristotelian Society* sup. vol. XXVIII: 63–76.

Bach, K. (1975) 'Performatives are statements too', *Philosophical Studies* vol. 28: 229–236.

Baker, G. P. and Hacker, P. M. S. (1984) *Scepticism, Rules and Language*, Oxford: Blackwell.

Barnes, B. (1977) *Interests and the Growth of Knowledge*, London: Routledge.

—— (1982) *T. S. Kuhn and Social Science*, London: Macmillan.

—— (1983) 'Social life as bootstrapped induction', *Sociology* vol. 4: 524–545.

—— (1988) *The Nature of Power*, Cambridge: Polity Press.

Bennett, J. (1961) 'On being forced to a conclusion', *Proceedings of the Aristotelian Society* sup. vol. XXXV: 15–34.

Blackburn, S. (1981) 'Rule following and moral realism', in S. H. Holtzman and C. M. Leich (eds) *Wittgenstein: To Follow a Rule*, London: Routledge: 163–187.

—— (1984) 'The individual strikes back', *Synthese* vol. 58: 281–301.

Bloor, D. (1973) 'Wittgenstein and Mannheim on the sociology of mathematics', *Studies in the History and Philosophy of Science* vol. 4: 173–191.

—— (1974) 'Popper's mystification of objective knowledge', *Science Studies* vol. 4: 65–76.

—— (1983) *Wittgenstein: A Social Theory of Knowledge*, London: Macmillan.

—— (1989) 'Wittgenstein on rule following: the old and the new individualism', *The Polish Sociological Bulletin*, no. 3–4: 27–33.

—— (1991) *Knowledge and Social Imagery*, Chicago: Chicago University Press.

—— (1992) 'Left and right Wittgensteinians', in A. Pickering (ed.) *Science as Practice and Culture*, Chicago: University of Chicago Press: 266–282.

—— (1994) 'What can the sociologist of knowledge say about 2 + 2 = 4 ?', in P. Ernest (ed.) *Mathematics, Education and Philosophy*, London: Falmer: Chapter 2.

—— (1996) 'The question of linguistic idealism revisited', in H. Sluga and D. Stern (eds) *The Cambridge Companion to Wittgenstein*, Cambridge: Cambridge University Press.

—— (1996a) 'Idealism and the sociology of knowledge', *Social Studies of Science*, (forthcoming).

—— (1997) 'Wittgenstein's behaviourism' in W. O'Donohue and R. Kitchener (eds) *Handbook of Behaviorism*, San Diego, Cal.: Academic Press (forthcoming).

Boghossian, P. A. (1989) 'The rule-following considerations', *Mind* vol. 98: 507–549.

Brannigan, A. (1981) *The Social Basis of Scientific Discoveries*, Cambridge: Cambridge University Press.

Budd, M. (1984) 'Wittgenstein on meaning, interpretation and rules', *Synthese* vol. 58: 759–778.

Carruthers, P. (1986) 'Review of McGinn (1984)', *Philosophical Books*, vol. 27: 36–38.

Champlin, T. S. (1992) 'Solitary rule-following', *Philosophy*, vol. 67: 285–306.

Chihara, C. (1982) 'The Wright-wing defence of Wittgenstein's philosophy of logic', *Philosophical Review*, vol. XCL: 99–108.

Churchland, P. M. (1989) *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, Mass.: MIT Press.

Coates, P. (1986) 'Kripke's sceptical paradox: normativeness and meaning', *Mind* vol. 95: 77–80.

Collins, A. (1992) 'On the paradox Kripke finds in Wittgenstein', in P. A. French, T. E. Uehling and H. K. Wettstein (eds) *The Wittgenstein Legacy. Midwest Studies in Philosophy* vol. XVII: 74–88, Notre Dame, Ind.: University of Notre Dame Press.

Dauben, J. W. (1979) *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Princeton NJ, Princeton University Press.

Davie, G. E. (1991) *The Scottish Enlightenment and Other Essays*, Edinburgh: Polygon.

—— (1994) *A Passion for Ideas: Essays on the Scottish Enlightenment*, Edinburgh: Polygon.

Defoe, D. ([1719] 1987) *Adventures of Robinson Crusoe*, Leicester: Galley Press.

Desmond, A. (1989) *The Politics of Evolution: Morphology, Medicine, and Reform in Radical London*, Chicago: Chicago University Press.

Dummett, M. (1959) 'Wittgenstein's Philosophy of Mathematics', *Philosophical Review* vol. 68: 324–348.

Durkheim, E. (1893) *The Division of Labour in Society*, trans. (1933) G. Simpson, London and New York: Collier Macmillan.

—— (1914) 'The dualism of human nature and its social conditions', in K. H. Wolf (ed.) (1964) *Essays on Sociology and Philosophy*, New York: Harper & Row: 325–340.

—— (1915) *The Elementary Forms of the Religious Life*, trans. J. W. Swain, New York: Collier.

Findlay, J. N. (1984) *Wittgenstein: a Critique*, London: Routledge.

Fodor, J. A. (1975) *The Language of Thought*, Cambridge, Mass.: Harvard University Press.

—— (1987) *Psychosemantics*, Cambridge, Mass.: MIT Press.

Forbes, G. (1984) 'Scepticism and semantic knowledge', *Proceedings of the Aristotelian Society*, vol. LXXXIV: 223–237.

Frege, G. (1884) *The Foundations of Arithmetic*, trans J. L. Austin (1959), Oxford: Blackwell.

French, P. A., Uehling, T. E., Wettstein, H. K. (eds) (1992) *The Wittgenstein Legacy. Midwest Studies in Philosophy*, vol. XVII, Notre Dame, Ind.: University of Notre Dame Press.

Gilbert, M. (1983) 'On the question of whether language has a social nature: some aspects of Winch and others on Wittgenstein', *Synthese*, vol. 56: 301–318.

Ginet, C. (1992) 'The dispositionalist solution to Wittgenstein's problem about understanding a rule: answering Kripke's objections', in P. A. French *et al.* (eds) *The Wittgenstein Legacy. Midwest Studies in Philosophy*, vol. XVII, Notre Dame, Ind.: University of Notre Dame Press.

Goldfarb, W. (1985) 'Kripke on Wittgenstein on rules', *Journal of Philosophy* vol. LXXXII: 471–488.

Harrison, B. (1991) 'Wittgenstein and Scepticism', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 34–69.

Haugeland, J. (1982) 'Weak supervenience', *American Philosophical Quarterly* vol. 19: 93–103.

—— (1990) 'The intentionality all-stars', in J. E. Tomberlin (ed.) *Philosophical Perspectives, 4, Action Theory and Philosophy of Mind*, Ataxadero, Cal.: Ridgeview: 383–427.

Hesse, M. (1974) *The Structure of Scientific Inference*, London: Macmillan.

Hill, C. (1975) *The World Turned Upside Down. Radical Ideas During the English Revolution*, Harmondsworth: Penguin.

Hoffman, P. (1985) 'Kripke on private languages', *Philosophical Studies* vol. 47: 23–28.

Hume, D. (1739–40) *A Treatise of Human Nature*, L. A. Selby-Bigge (ed.) (1960), Oxford: Clarendon.

Keynes, J. M. (1973) *The Collected Writings*, vol. III, London: Macmillan.

Kitcher, P. (1984) *The Nature of Mathematical Knowledge*, Oxford: Oxford University Press.

Kripke, S. (1982) *Wittgenstein on Rules and Private Language*, Oxford: Blackwell.

Krishna, D. (1971) 'The self-fulfilling prophecy and the nature of society', *American Sociological Review*, vol. 36: 1104–1107.

Kuhn, T. S. (1962) 'The historical structure of scientific discovery', *Science* vol. 136: 760–764.

—— (1963) 'The function of dogma in scientific research', in A. C. Crombie (ed.) *Scientific Change*, London: Heinemann: 349–369 and 386–395.

Lewis, D. K. (1969) *Convention: A Philosophical Study*, Oxford: Blackwell.

Locke, J. (1699) *An Essay Concerning Human Understanding*, A. S. Pringle-Patterson (ed.) (1924), Oxford: Clarendon.

Loar, B. (1985) 'Critical review of Kripke (1982)', *Nous*, vol. XIX: 275–280.

Luntley, M. (1991) 'The transcendental grounds of meaning and the place of silence', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 170–188.

McDonough, R. (1991) 'Wittgenstein's refutation of meaning scepticism', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 70–92.

McDowell, J. (1981) 'Non-cognitivism and rule-following', in S. H. Holtzman and C. M. Leich (eds) *Wittgenstein: To Follow a Rule*, London: Routledge: 141–162.

—— (1984) 'Wittgenstein on following a rule', *Synthese* vol. 58: 325–363.

—— (1991) 'Intentionality and interiority in Wittgenstein', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 148–169.

—— (1992) 'Meaning and intentionality in Wittgenstein's later philosophy' in P. A. French, T. E. Uehling and H. K. Wettstein (eds) *The Wittgenstein Legacy. Midwest Studies in Philosophy* vol. XVII: 40–52, Notre Dame, Ind.: University of Notre Dame Press.

McGinn, C. (1984) *Wittgenstein on Meaning: An Interpretation and Evaluation*, Oxford: Blackwell.

—— (1989) *Mental Content*, Oxford: Blackwell.

MacIntyre, A. (1967) *A Short History of Ethics*, London: Routledge.

—— (1988) *Whose Justice? Which Rationality?*, London: Duckworth.

Mackie, J. L. (1966) 'Proof', *Proceedings of the Aristotelian Society*, sup. vol. XL: 23–38.

—— (1980) *Hume's Moral Theory*, London: Routledge.

McKinley, A. (1991) 'Agreement and normativity', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 189–200.

Malcolm, N. (1954) 'Wittgenstein's "Philosophical Investigations"', *Philosophical Review* vol. LXIII: 530–559.

—— (1958) *Ludwig Wittgenstein: A Memoir*, London: Oxford University Press.
—— (1982) 'Wittgenstein: the relation of language to instinctive behaviour', *Philosophical Investigations* vol. 5: 3–22.
—— (1989) 'Wittgenstein on language and rules', *Philosophy*, vol. 64: 5–28.
—— (1993) *Wittgenstein: A Religious Point of View?*, London: Routledge.
Mandelbaum, M. (1955) 'Societal facts', *British Journal of Sociology* vol. VI: 305–317.
Mannheim, K. (1936) *Ideology and Utopia: An Introduction to the Sociology of Knowledge*, London: Routledge.
Margalit, A. (1992) 'How to outsmart the rules. A comment,' (A reply to Smart 1992), in E. Ullmann-Margalit (ed.) *The Scientific Enterprise*, Dordrecht: Kluwer.
Menger, C. (1892) 'On the origin of money', trans. C. A. Foley, *Economic Journal* vol. 2: 239–255 (reprinted in S. Littlechild (ed.) (1990) *Austrian Economics* vol. II, London: Elgar: Chapter 9).
Merton, R. K. (1968) *Social Theory and Social Structure*, (third edn), New York: Free Press.
Mill, J. S. (1848) *A System of Logic: Ratiocinative and Inductive*, London: Longmans (impression of eighth edn 1959).
Millikan, R. G. (1984) *Language, Thought and Other Biological Categories*, Cambridge, Mass.: MIT Press.
—— (1990) 'Truth rules, hoverflies, and the Kripke–Wittgenstein paradox', *Philosophical Review* vol. XCIX: 323–353.
Mises, L. von (1935) 'Economic calculation in the socialist commonwealth', in F. A. Hayek (ed.) *Collectivist Economic Planning*, London: Routledge: 87–130.
—— (1963) *Human Action: A Treatise on Economics*, (third rev. edn), Chicago: Contemporary Books.
Moore, G. H. (1982) *Zermelo's Axiom of Choice. Its Origins, Development and Influence*, Heidelberg: Springer.
Murray, G. (1911) *Oedipus King of Thebes by Sophocles*, London: Allen & Unwin.
Neurath, O. (1973) *Empiricism and Sociology*, M. Neurath and R. S. Cohen (eds), Dordrecht: Reidel: Chapter 5.
Olby, R. (1979) 'Mendel no Mendelian?', *History of Science* vol. 17: 53–72.
Olson, M. (1965) *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, Mass.: Harvard University Press.
Paton, H. J. (1946) *The Categorical Imperative*, London: Hutchinson.
Peacocke, C. (1981) 'Rule-following: the nature of Wittgenstein's arguments', in S. H. Holtzman and C. M. Leich (eds) *Wittgenstein: to Follow a Rule*, London: Routledge: Chapter II.
Pettit, P. (1990) 'The reality of rule-following', *Mind* vol. 99: 1–21.
Popper, K. R. (1969) *The Open Society and its Enemies* vol. II (fifth rev. edn), London: Routledge.
—— (1972) *Objective Knowledge; An Evolutionary Approach*, Oxford: Clarendon.
Prior, A. N. (1949) *Logic and the Basis of Ethics*, Oxford: Clarendon.

Puhl, K. (ed.) (1991) *Meaning Scepticism*, Berlin: Walter de Gruyter.
Reid, T. (1785) *Essays on the Intellectual Powers of Man* (ed. with notes, etc.) G. N. Wright (1843), London: Thomas Tegg.
—— (1788) *Essays on the Active Powers of the Human Mind*, introduction by B. A. Brody (1969), Cambridge, Mass.: MIT Press.
Rhees, R. (1954) 'Can there be a private language?', *Proceedings of the Aristotelian Society*, sup. vol. XXVII: 77–94.
Rudwick, M. J. S. (1972) *The Meaning of Fossils: Episodes in the History of Palaeontology*, London: MacDonald.
Russell, B. (1919) *Introduction to Mathematical Philosophy*, London: Allen & Unwin.
Ryle, G. (1949) *The Concept of Mind*, London: Hutchinson.
Sartorelli, J. (1991) 'Review of Baker and Hacker (1984)', *Philosophical Review* vol. 100: 660–662.
Schelling, T. C. (1960) *The Strategy of Conflict*, Cambridge, Mass.: Harvard University Press.
Schiffer, S. (1986) 'Kripkenstein meets the remnants of meaning', *Philosophical Studies* vol. 49: 147–162.
Schlick, M. (1939) *Problems of Ethics*, trans. D. Rynin (1962), New York: Dover.
Searle, J (1995) *The Social Construction of Reality*, Harmondsworth: Penguin.
Shanker, S. G. (1984) 'Sceptical confusions about rule-following', *Mind* vol. 93: 423–429.
—— (1988) 'Wittgenstein's remarks on the significance of Gödel's theorem', in S. G. Shanker (ed.) *Gödel's Theorem in Focus*, London: Croom Helm: 155–256.
Shapin, S. (1991) '"The mind is its own place": science and solitude in seventeenth-century England', *Science in Context* vol. 4: 191–218.
—— (1994) *A Social History of Truth: Civility and Science in Seventeenth-Century England*, Chicago: Chicago University Press.
—— and Schaffer, S. (1985) *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*, Princeton, NJ: Princeton University Press.
Smart, J. J. C. (1992) 'Wittgenstein, following a rule, and scientific psychology', in E. Ullmann-Margalit (ed.) *The Scientific Enterprise*, Dordrecht: Kluwer: 123–137.
Stabler, E. P. (1987) 'Kripke on functionalism and automata', *Synthese* vol. 70: 1–22.
Steele, D. R. (1992) *From Marx to Mises: Post-Capitalist Society and the Challenge of Economic Calculation*, La Salle, Ill.: Open Court.
Strawson, P. F. (1954) 'Review of Wittgenstein's *Philosophical Investigations*', *Mind* vol. LXIII: 70–99.
Stroud, B. (1965) 'Wittgenstein and logical necessity', *The Philosophical Review* vol. 74: 504–518.
Summerfield, D. M. (1990) '"Philosophical Investigations" 201: A Wittgensteinian reply to Kripke', *Journal of the History of Philosophy* vol. 28: 417–438.

Tönnies, F. (1887) *Community and Society (Gemeinschaft u. Gesellschaft)*, trans. C. P. Loomis (1955), London: Routledge.

Toulmin, S. E. (1965) *The Philosophy of Science: An Introduction*, London: Hutchinson.

Waismann, F. (1965) 'Notes on talks with Wittgenstein', *Philosophical Review* vol. 74: 12–16.

—— (1979) *Wittgenstein and the Vienna Circle*, B. McGuiness (ed.), Oxford: Blackwell.

Werhane, P. H. (1987) 'Some paradoxes in Kripke's interpretation of Wittgenstein', *Synthese* vol. 73: 253–273.

Will, F. L. (1985) 'Rules and subsumption: mutative aspects of logical processes', *American Philosophical Quarterly*, vol. 22: 143–151.

Williams, M. (1983) 'Wittgenstein on representation, privileged objects, and private languages', *Canadian Journal of Philosophy* vol. XIII: 57–78.

—— (1985) 'Wittgenstein's rejection of scientific psychology', *Journal for the Theory of Social Behaviour* vol. 15: 203–223.

—— (1990) 'Social norms and narrow content', in P. A. French, T. E. Uehling and H. K. Wettstein (eds) *The Wittgenstein Legacy. Midwest Studies in Philosophy* vol. XVII: 425–462, Notre Dame, Ind.: University of Notre Dame Press.

—— (1991) 'Blind obedience: rules, community and the individual', in K. Puhl (ed.) *Meaning Scepticism*, Berlin: Walter de Gruyter: 93–125.

—— (1994a) 'The significance of learning in Wittgenstein's later philosophy', *Canadian Journal of Philosophy* vol. 24: 173–204.

—— (1994 b) 'Private states and public practices: Wittgenstein and Schutz on intentionality', *International Philosophical Quarterly* vol. XXXIV: 89–110.

Winch, P. (1958) *The Idea of a Social Science and its Relation to Philosophy*, London: Routledge.

Winsor, M. P. (1976) *Starfish, Jellyfish and the Order of Life*, New Haven, Conn.: Yale University Press.

Wittgenstein, L. (1967) *Philosophical Investigations*, trans. G. E. M. Anscombe, Oxford: Blackwell.

—— (1967) *Zettel*, G. E. M. Anscombe and G. H. von Wright (eds), trans. G. E. M. Anscombe, Oxford: Blackwell.

—— (1969) *The Blue and Brown Books*, Oxford: Blackwell.

—— (1969) *On Certainty*, G. E. M. Anscombe and G. H. von Wright (eds), trans. D. Paul and G. E. M. Anscombe, Oxford: Blackwell.

—— (1978) *Remarks on the Foundations of Mathematics*, G. H. von Wright, R. Rhees and G. E. M. Anscombe (eds), trans. G. E. M. Anscombe, Oxford: Blackwell.

—— (1980) *Culture and Value*, G. H. von Wright and H. Nyman (eds), trans. P. Winch, Oxford: Blackwell.

Wright, C. (1984) 'Kripke's account of the argument against private languages', *Journal of Philosophy* vol. 81: 759–778.

—— (1986) 'Rule-following, meaning and construction', in C. Travis (ed.) *Meaning and Interpretation*, Oxford: Blackwell.

NAME INDEX

SUBJECT INDEX